



# Détection de changements et classification sous-pixeliques en imagerie satellitaire. Application au suivi temporel des surfaces continentales.

Amandine Robin

## ► To cite this version:

Amandine Robin. Détection de changements et classification sous-pixeliques en imagerie satellitaire. Application au suivi temporel des surfaces continentales.. Mathématiques [math]. Université René Descartes - Paris V, 2007. Français. NNT: . tel-00163361

**HAL Id: tel-00163361**

**<https://theses.hal.science/tel-00163361>**

Submitted on 17 Jul 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNIVERSITÉ PARIS DESCARTES**

UFR de Mathématiques et Informatique

École doctorale Paris-Centre

**THÈSE**

pour obtenir le grade de

**Docteur de l'Université Paris Descartes**

Spécialité : **Mathématiques appliquées**

présentée par

**Amandine ROBIN**

---

**Détection de changements et classification  
sous-pixeliques en imagerie satellitaire.**

**Application au suivi temporel des surfaces continentales.**

---

**Sub-pixelic change detection and classification.**

**Application to remote-sensing land-cover monitoring.**

---

Soutenue publiquement le 21 Mai 2007 devant le jury composé de :

Président :	<b>Jean-Michel MOREL</b>	Professeur, École Normale Supérieure de Cachan
Directeurs de thèse :	<b>Sylvie LE HÉGARAT-MASCLE</b> <b>Lionel MOISAN</b>	Professeur, Université Paris Sud 11 Professeur, Université de Paris Descartes
Rapporteurs :	<b>Xavier DESCOMBES</b> <b>Sebastiano SERPICO</b>	Chargé de Recherche INRIA Professeur, Université de Gènes, Italie
Examineurs :	<b>Hervé CARDOT</b> <b>Jordi INGLADA</b> <b>Hervé POILVÉ</b>	Chargé de Recherche INRA Ingénieur de Recherche, CNES Ingénieur de Recherche, ASTRIUM/EADS



## Remerciements

Cette page est dédiée à tous ceux qui m'ont soutenue, encouragée ou éclairée pendant ces quatre dernières années. Je souhaite leur exprimer ici toute ma reconnaissance.

En premier lieu, je voudrais exprimer toute ma gratitude à Sylvie Le Hégarat-Masclé et à Lionel Moisan qui ont accepté de co-encadrer ce travail. Je n'aurais pas pu le mener à bien sans leur dynamisme, leur exigence et leur patience. Leurs qualités scientifiques et humaines, très complémentaires, m'ont beaucoup apporté.

Je tiens à remercier tout spécialement Jordi Inglada, non seulement pour sa participation au jury, mais surtout pour son accompagnement tout au long de ma thèse. Sa disponibilité et son regard différent et pertinent sur mes travaux a apporté, régulièrement, un nouveau souffle dans ce travail. Je lui suis aussi très reconnaissante pour m'avoir permis, si facilement, d'exploiter des images SPOT.

Je voudrais également remercier Hervé Poilvé et son équipe pour l'intérêt qu'ils ont porté à ce travail. En m'invitant à passer deux mois au sein du département Recherche et Développement "Agriculture et Environnement" (Infoterra, EADS Astrium) au début de ma thèse, ils m'ont permis de découvrir une discipline qui m'était alors parfaitement inconnue : l'agriculture de précision.

Xavier Descombes et Sebastiano Serpico ont accepté d'être les rapporteurs de cette thèse. Je les remercie chaleureusement pour l'attention avec laquelle ils ont lu ce document. Leurs remarques judicieuses sont autant de pistes de recherche à explorer.

Je suis très reconnaissante à Hervé Cardot d'avoir accepté de faire partie de mon jury. Je le remercie pour l'intérêt qu'il a porté à ce travail et pour ses remarques constructives lors de la soutenance.

Je remercie Jean-Michel Morel de m'avoir fait l'honneur de présider mon jury de thèse. Je lui suis également très reconnaissante pour sa disponibilité, sa bienveillance et ses conseils lors de ma première année de thèse qui m'ont permis de bénéficier d'un excellent cadre de travail.

Mes remerciements vont également à Antoine Chambaz pour son enthousiasme remarquable et pour m'avoir éclairé sur un certain nombre de questions statistiques, à Yves Rozenholc pour ses réflexions fructueuses et sa grande disponibilité, ainsi qu'Alain Trouvé pour m'avoir fait profiter de sa grande expérience en matière de recuit simulé.

Durant ces quatre années, j'ai bénéficié au quotidien d'excellentes conditions de travail au sein de trois laboratoires : le Centre d'étude des Environnements Terrestres et Planétaires (CETP), le Centre de Mathématiques et Leurs Applications (CMLA) et le laboratoire de Mathématiques Appliquées de Paris 5 (MAP5). Tous trois m'ont offert un accueil sympathique et convivial ainsi qu'un cadre de travail stimulant. J'en remercie chaleureusement tous les membres.

Je ne serais sans doute pas arrivée jusqu'ici sans avoir suivi les cours d'enseignants exceptionnels. Je pense en particulier à Carmen Mafté en terminale et à tous les enseignants de Paris 5 que j'ai eu le plaisir de retrouver ces dernières années. Je dois également des remerciements particuliers à Georges Koepfler pour sa disponibilité et sa bienveillance et à Agnès Desolneux notamment pour ses encouragements dans la dernière ligne droite.

Pour les questions plus pratiques ou administratives, j'adresse des remerciements spéciaux à Nellie Bouchard, Christophe Castellani (et son aide précieuse pour les impressions couleurs...), Marie-Hélène Gbaguidi, Voëuni Kheng ainsi qu'à Armande Zénarre. Leur disponibilité, leur bonne humeur et leur efficacité ont été particulièrement appréciables. Je salue également Micheline et Véro pour leur énergie infaillible et communicative. Je dois également des remerciements particuliers aux ingénieurs informaticiens

du CMLA et de Paris 5 : Pascal Bringas, Christophe Labourdette, Azedine Mani, Laurent Moineau et Thierry Raedersdoff. Leur disponibilité et leur savoir-faire ont été très précieux pour réparer mes nombreuses *boulettes*.

Je voudrais aussi souligner l'influence positive de toutes les personnes dont j'ai partagé le quotidien ces dernières années. Dans un ordre chronologique, je pense à Julie et Ben pour leur présence et leur gentillesse hors norme. Je pense également à toute la fine équipe des thésards du CMLA et aux mémorables parties de tennis-bouteille sur les belles pelouses de l'ENS, sans oublier Papy et ses célèbres tea parties !

Papy, merci pour ta présence et ton réconfort tout au long de ces dernières années, mais aussi pour m'avoir fait profiter de ton esprit critique magnifiquement développé tout en m'accompagnant dans la découverte des meilleurs cocktails *happy hour* de Saint-Germain-des-Près !

Je ne saurais oublier l'ensemble des doctorants et jeunes docteurs du MAP5 (d'aujourd'hui et d'hier) avec lesquels j'ai finalement passé la plupart de mon temps ! Merci à Béatrice, Javiera et Olivier et leur exceptionnelle créativité pour alimenter les pauses déjeuner. Qu'aurions-nous fait sans e-pousse, ouverture-facile et les volcans de chocolat ? Je pense également au gang des recycleurs de sachets de thé et receleurs de chocolat allemand : Arno, Cécile, Claire J., Claire L. (ma plus proche voisine et tout ce que ça implique), Gwendo, Nicolas et Sylvain, mes compagnons de bureau. Leur présence m'a beaucoup apporté, en particulier dans la dernière ligne droite. Je salue Arno pour me faire le plaisir de toujours perdre ses paris, au profit des merveilleux beignets du coin. Je n'oublie pas non plus Hermine dont les encouragements me furent précieux, ni Béné et ses rassurants debriefing.

Un grand merci à Claire et tous ces bons moments partagés depuis la licence, sans oublier de saluer l'incontestable efficacité du CSFDAR ! Je remercie aussi l'irremplaçable oncle Jaco.

Merci à Benjamin pour les irréfutables "points sur la thèse" et surtout pour les divagations éthyliques de la dernière chance ! Je ne saurais oublier ce pari perdu et ses délicates incitations à sldc.

Une mention spéciale à Aïcha et Jacopo pour leur soutien continu ces dernières années et leur indéfectible enthousiasme, sans oublier leurs hautes performances respectives en tennis et en squash.

J'en profite pour rendre hommage à toutes les séries TV qui m'ont guidée spirituellement ces dernières années... Je ne peux que remercier platement les dealers plus ou moins licites qui m'ont approvisionnée.

Il me reste à remercier chaleureusement ma famille pour m'avoir supportée toutes ces années, et en particulier Seved pour m'avoir patiemment transmis le goût des maths.

Enfin, merci à Juju pour m'avoir choyée, épaulée et rassurée depuis le début de cette aventure. Du fond du coeur, merci.





# Table des matières

Notations . . . . .	6
<b>Introduction générale</b>	<b>9</b>
<b>I. Contexte de l'étude et objectifs</b>	<b>15</b>
<b>1. Acquisitions satellitaires et observation de la surface terrestre</b>	<b>17</b>
1.1. Introduction . . . . .	17
1.2. Les capteurs . . . . .	18
1.2.1. Caractéristiques des systèmes imageurs . . . . .	18
1.2.2. Les satellites d'observation de la Terre . . . . .	21
1.3. Spécificités liées à la végétation . . . . .	25
1.3.1. Les propriétés spectrales . . . . .	25
1.3.2. Les indices de végétation . . . . .	27
1.3.3. L'aspect temporel . . . . .	30
1.4. Conclusion . . . . .	31
<b>2. Méthodes de détection de changements</b>	<b>33</b>
2.1. Introduction . . . . .	33
2.1.1. Problématiques de la détection de changements . . . . .	33
2.1.2. Taxonomie des méthodes de détection de changements . . . . .	34
2.2. Prétraitements . . . . .	36
2.3. Les approches ponctuelles . . . . .	37
2.3.1. Les transformations . . . . .	38
2.3.2. La décision . . . . .	41
2.3.3. Les méthodes à base de classifications . . . . .	43
2.4. Les approches contextuelles . . . . .	45
2.4.1. Les transformations . . . . .	46
2.4.2. La décision . . . . .	47
2.4.3. Les méthodes à base de classifications . . . . .	48
2.5. Conclusion . . . . .	48



<b>3. Méthodes d'analyse sous-pixellique</b>	<b>51</b>
3.1. Motivation . . . . .	51
3.2. Les méthodes de fusion de données multi-résolutions . . . . .	52
3.3. La désagrégation de l'information . . . . .	54
3.4. La classification sous-pixellique . . . . .	55
3.5. Conclusion . . . . .	57
<b>4. Problématique</b>	<b>59</b>
 <b>II. Classification sous-pixelique de l'occupation du sol</b>	 <b>61</b>
<b>5. Un modèle bayésien pour la classification</b>	<b>65</b>
5.1. Modélisation du problème . . . . .	65
5.2. Maximum <i>a posteriori</i> . . . . .	69
5.3. Influence de la basse résolution . . . . .	74
5.4. Analyse d'erreurs et aspects multitemporels . . . . .	78
<b>6. Algorithmes et performances empiriques</b>	<b>85</b>
6.1. Algorithme de recuit simulé . . . . .	85
6.2. Algorithme de programmation dynamique . . . . .	87
6.3. Simulation des données . . . . .	90
6.4. Exemples d'application . . . . .	91
6.5. Analyse des erreurs . . . . .	95
6.6. Sensibilité au facteur de résolution . . . . .	97
 <b>III. Détection de changements à l'échelle sous-pixelique</b>	 <b>101</b>
<b>7. Un modèle <i>a contrario</i> pour la détection de changements</b>	<b>103</b>
7.1. Modélisation du problème . . . . .	103
7.2. Détection <i>a contrario</i> . . . . .	105
7.3. Choix du modèle <i>a contrario</i> . . . . .	114
7.4. Modèle multidimensionnel . . . . .	120
7.5. Comparaison avec des procédures statistiques classiques . . . . .	122
7.5.1. Les procédures de tests multiples . . . . .	122
7.5.2. L'analyse de variance . . . . .	125
<b>8. Aspects numériques</b>	<b>129</b>
8.1. Problématique . . . . .	129
8.2. Algorithme . . . . .	131
8.2.1. Cas monodimensionnel . . . . .	131
8.2.2. Cas multidimensionnel . . . . .	134

8.2.3. Exemple d'application monotemporelle . . . . .	136
8.3. Nombre d'itérations nécessaire . . . . .	138
8.4. Sélection de sous-systèmes . . . . .	142
<b>9. Performances théoriques</b>	<b>153</b>
9.1. Modélisation des données . . . . .	153
9.2. Sensibilité au niveau de contraste . . . . .	154
9.3. Influence de la taille des images . . . . .	157
9.4. Etude en fonction du facteur de résolution . . . . .	161
<b>10. Performances empiriques</b>	<b>165</b>
10.1. Evaluation de la performance des résultats . . . . .	165
10.2. Simulation des données sans changements . . . . .	167
10.3. Influence du nombre de pixels sans changement . . . . .	168
10.3.1. Simulation des changements . . . . .	169
10.3.2. Résultats . . . . .	170
10.4. Influence de la taille du changement sous-pixellique . . . . .	177
10.4.1. Simulation des changements . . . . .	177
10.4.2. Résultats . . . . .	178
<b>IV. Application à une zone agricole de la plaine du Danube (base de données ADAM)</b>	<b>183</b>
<b>11. Introduction</b>	<b>185</b>
11.1. La base de données ADAM . . . . .	185
11.2. Quelques scenarii d'application . . . . .	188
<b>12. Classification et validation</b>	<b>191</b>
12.1. Classification sous-pixellique . . . . .	191
12.2. Validation d'une carte d'occupation des sols . . . . .	198
12.3. Choix du nombre de classes . . . . .	205
<b>13. Détection de changements et mise à jour de classification</b>	<b>211</b>
13.1. Détection de changements . . . . .	211
13.2. Mise à jour de classification . . . . .	213
<b>Conclusion générale</b>	<b>219</b>
Publications . . . . .	223
<b>Bibliographie</b>	<b>227</b>

<b>Liste des figures</b>	<b>235</b>
--------------------------	------------



# Notations

## Symboles mathématiques - Symboles courants

$\mathcal{D}$	: le domaine géographique (rectangulaire) observé.
$\mathcal{D}_{\text{HR}}$	: le support discret de l'image haute résolution, $\mathcal{D}_{\text{HR}} = \{1, \dots, m\} \times \{1, \dots, n\}$ la grille de pixels haute résolution qui découpe le domaine $\mathcal{D}$ en $m \times n$ pixels.
$x$	: un pixel du domaine $\mathcal{D}_{\text{HR}}$ .
$\mathcal{D}_{\text{BR}}$	: le support discret de l'image basse résolution, $\mathcal{D}_{\text{BR}} = \{1, \dots, m'\} \times \{1, \dots, n'\}$ la grille de pixels basse résolution qui découpe le domaine $\mathcal{D}$ en $m' \times n'$ pixels.
$y$	: un pixel du domaine $\mathcal{D}_{\text{BR}}$ .
$\{x \in y\}$	: l'ensemble des pixels $x \in \mathcal{D}_{\text{HR}}$ représentant une partie de la scène couverte aussi par le pixel $y \in \mathcal{D}_{\text{BR}}$ .
$ \cdot $	: le cardinal d'un ensemble.
$N$	: le rapport entre les résolutions considérées, $N =  \mathcal{D}_{\text{HR}} / \mathcal{D}_{\text{BR}} $ . Remarque : $ W_y  = N$ .
$\mathcal{T}$	: l'ensemble des différentes dates d'acquisitions disponibles, $\mathcal{T} = \{1, \dots, T\}$ .
$u$	: une séquence d'images haute résolution, $u = (u_1, \dots, u_T)$ .
$u_t$	: l'image associée à la date $t$ de la séquence haute résolution ( $t \in \mathcal{T}$ ),

$$\begin{aligned} u_t : \mathcal{D}_{\text{HR}} &\rightarrow \mathbb{R} \\ x &\mapsto u_t(x). \end{aligned}$$

$U$	: le champ aléatoire dont $u$ est une réalisation.
$U_t(x)$	: la variable aléatoire réelle correspondante dont l'intensité mesurée $u_t(x)$ est une réalisation.
$v$	: une séquence d'images basse résolution, $v = (v_1, \dots, v_T)$ .
$v_t$	: l'image associée à la date $t$ de la séquence basse résolution ( $t \in \mathcal{T}$ ),

$$\begin{aligned} v_t : \mathcal{D}_{\text{BR}} &\rightarrow \mathbb{R} \\ y &\mapsto v_t(y). \end{aligned}$$

$V$	: le champ aléatoire dont $v$ est une réalisation.
-----	--

- $\mathcal{S}$  : l'ensemble des segments  $\mathcal{S} = \{1, \dots, S\}$  d'une segmentation.  
 $s$  : une segmentation du domaine  $\mathcal{D}_{\text{HR}}$ , définie par l'application

$$\begin{aligned} s : \mathcal{D}_{\text{HR}} &\rightarrow \mathcal{S} = \{1, \dots, S\} \\ x &\mapsto s(x). \end{aligned}$$

- $\mathcal{L}$  : l'ensemble des labels  $\mathcal{L} = \{1, \dots, L\}$  d'une labellisation (correspondant aux différents types d'occupation du sol).  
 $\lambda$  : une labellisation, notée  $\lambda = (\lambda_1, \dots, \lambda_S)$ , est définie par l'application

$$\begin{aligned} \lambda : \mathcal{S} &\rightarrow \mathcal{L} = \{1, \dots, L\} \\ k &\mapsto \lambda_k, \end{aligned}$$

où  $\lambda_k$  représente le label de la région  $k$ .

- $\Lambda$  : le champ aléatoire discret dont  $\lambda$  est une réalisation. Il est défini sur  $\mathcal{S}$ .  
 $\mathcal{M}(n, p)$  : l'ensemble des matrices de taille  $n \times p$  à coefficients réels.  
 $L^2(\mathbb{R})$  : l'espace des fonctions de carré intégrable sur  $\mathbb{R}$ .  
 $\|\cdot\|_2$  : la norme sur  $L^2(\mathbb{R})$ .  
 $\mathbb{E}[\cdot]$  : une espérance.  
 $\mathbb{P}(\cdot)$  : une distribution de probabilité.

## Principaux acronymes et abréviations

- BR** : une image Basse Résolution.  
**HR** : une image Haute Résolution.  
**R** : mesure dans le canal Rouge.  
**PIR** : mesure dans le canal proche-infrarouge.  
**VIS** : la mesure dans le canal VISible.  
**MAP** : Maximum A Posteriori.  
**SPOT** : Satellite Pour l'Observation de la Terre.  
**HRVIR** : Haute Résolution Visible et Infra-Rouge.  
**VGT** : Végétation.  
**MeRIS** : Medium Resolution Imaging Spectrometer.  
**NOAA** : National Oceanic and Atmospheric Administration.  
**AVHRR** : Advanced Very High Resolution Radiometer.



# Introduction générale

La recherche effectuée dans le cadre de cette thèse porte sur le suivi de l'évolution des surfaces continentales et, en particulier, de la végétation. Dans ce contexte, elle vise à l'élaboration d'outils automatiques de classification et de détection de changements à partir d'images acquises avec une haute fréquence temporelle et une basse résolution spatiale.

## Les enjeux

Les changements qui apparaissent sur les surfaces continentales peuvent être de nature très différente et provenir de phénomènes brutaux ou lents. Les changements ponctuels qui surviennent de manière brutale tels que les incendies, les inondations ou les éruptions volcaniques, sont généralement lourds de conséquences tant sur le plan humain que sur l'écosystème. Leur détection est alors essentielle pour la gestion des risques et des situations de crise. D'autres types de changements se manifestent sur des périodes plus longues (décadaires, annuelles ou décennales) de manière continue. Leur détection est alors nécessaire dans le cadre des processus de suivi et de contrôle de la surface, aussi bien pour les applications de géographie ou de biophysique que pour les applications liées aux interactions entre la surface et l'atmosphère et pour participer à la protection de l'environnement.

Les besoins en cartographie sont importants aujourd'hui tant pour la création de cartes, puisque plus de la moitié des terres émergées n'ont pas encore été cartographiées, que pour leur mise à jour puisque la surface continentale est en constante évolution (constructions, végétation, etc). Cette nécessité place l'élaboration de cartes et leur mise à jour au centre des préoccupations nationales et internationales, avec la création, notamment, du projet européen "CORINE land cover" initié en 1985 dont le but est de fournir une cartographie de référence sur l'ensemble de l'Europe. Ce projet vise à l'élaboration d'un inventaire biophysique de l'occupation des terres au 1/100 000ème ainsi qu'à leur mise à jour avec une fréquence décennale. Si la cartographie constitue un but en soi pour les applications géographiques, elle est aussi utilisée dans de nombreux domaines.

En effet, d'autres thématiques liées à l'étude du fonctionnement et de la dynamique des processus physiques ou à la compréhension de la mesure physique ne s'intéressent pas directement à la détection des changements mais nécessitent de connaître l'occupation des sols. Certains paramètres de surface sont à la base de l'inversion de modèles, notamment d'interactions entre les ondes et la surface (par exemple, les modèles de transfert radiatif) ou nécessaires pour comprendre la mesure acquise par des instruments de télédétection. De plus, à une échelle globale, les surfaces



continentales jouent un rôle fondamental dans les bilans énergétique, hydrique et carboné de la Terre. Elles interagissent étroitement avec l'atmosphère et le système hydrologique et évoluent constamment sous l'influence du climat et de l'action de l'homme.

Tout d'abord, l'étude de l'évolution des types de couvert est important pour la planification agricole (ex : la Politique Agricole Commune dans le cadre de l'Union Européenne) ainsi que pour l'agriculture de précision. En effet, le suivi de l'évolution des cultures à une échelle locale permet, par un dépistage précoce des maladies de la végétation et par un emploi adapté des engrais, l'optimisation de la production et le contrôle du rejet de substances polluantes (excédents de nitrates, etc.). À une échelle régionale, le suivi des zones agricoles est important pour l'environnement : par exemple, les changements des systèmes cultureux et des pratiques agricoles en Bretagne ont contribué, ces dernières décennies, à une dégradation importante des ressources en eau.

De plus, le suivi de la végétation permet de mieux appréhender l'activité photosynthétique et respiratoire des plantes pour en évaluer la production de biomasse ainsi que pour faire l'inventaire des sources et puits de carbone. Le contrôle des activités d'afforestation, de reforestation et de déforestation (ARD) est alors fondamental non seulement pour la gestion des forêts mais aussi pour le contrôle des gaz à effet de serre et, par conséquent, pour le respect des accords de Marrakech (Protocole de Kyoto). Ainsi, la surveillance du couvert végétal et de l'atmosphère constitue l'un des objectifs principaux du projet européen GMES<sup>1</sup> (Global Monitoring of Environment and Security). Elle permet, notamment, de contrôler précisément l'application des accords internationaux définis dans le cadre du changement climatique tels que le Protocole de Kyoto.

Enfin, les études liées au fonctionnement de la biosphère, au déterminisme des flux d'eau et de carbone ou aux processus atmosphériques et hydrologiques nécessitent des informations sur la surface. Les fortes relations existant entre les processus de surface et le système atmosphérique ont, notamment, donné lieu à de nombreux programmes dont l'objectif est de déterminer les échanges d'énergie et de masse (vapeur d'eau,  $CO_2$ , gaz divers, aérosols, etc.) entre la surface et l'atmosphère et de quantifier l'impact des variations climatiques sur ces échanges ainsi que sur le fonctionnement de la biosphère. Par exemple, l'objectif de l'IGBP (International Geosphere-Biosphere Program) est l'étude des changements globaux à travers l'analyse des processus biologiques, physiques et chimiques au cœur de la dynamique du système terrestre, l'analyse des changements ainsi que le rôle de l'activité humaine sur ces changements.

Le développement d'outils automatiques de classification et de détection de changements qui soient adaptés au suivi de la surface terrestre est alors nécessaire pour contribuer à la mise au point d'un système d'information permettant à la fois la mise à jour de cartographies et le contrôle de l'application des politiques communes, en particulier dans les domaines agricoles et

---

<sup>1</sup>Projet à l'initiative de l'ESA et de l'Union Européenne dans le but de fédérer les activités d'observation de la Terre et en vue de créer une capacité de surveillance à échelle locale, régionale et globale pour l'environnement et la sécurité en tant que support des politiques environnementales et agricoles européennes et des engagements internationaux de l'Union Européenne cf. <http://www.gmes.info/>

environnementaux.

## Le traitement d'images pour le suivi des surfaces continentales

La télédétection spatiale permet d'acquérir une quantité importante d'informations sur les surfaces continentales, à différentes échelles spatiales et temporelles. Leur utilisation dans les domaines biophysique ou atmosphérique a prouvé son efficacité, et les satellites construits sont de plus en plus performants et ciblés pour apporter des informations fines au service des applications visées.

Compte tenu des nombreux avantages offerts par l'imagerie satellitaire, le développement d'outils de traitement d'images et de modèles physiques surface-atmosphère a connu un essor marqué dans des domaines très divers durant ces dernières décennies. Les travaux réalisés peuvent être regroupés en deux grandes catégories : les approches dont l'objectif est d'interpréter physiquement le signal acquis à distance et celles qui cherchent à exploiter les propriétés intrinsèques des images pour leur traitement. Dans l'objectif d'interpréter la mesure physique acquise par télédétection, de nombreux modèles ont été créés pour prédire le signal des capteurs et, à l'inverse, pour estimer des paramètres de surface à partir des images (paramètres liés, par exemple, à la rugosité, à l'humidité ou à la végétation). De plus, le traitement d'images et la vision par ordinateur ont donné lieu à de nombreux modèles mathématiques pour la compréhension d'une scène ou d'un phénomène à partir d'images. Cependant, la quantité d'informations acquises par télédétection sur une base quotidienne ou mensuelle reste largement sous-exploitée. Le développement d'outils de traitement automatique des images est donc primordial pour l'exploitation de la masse des données disponibles.

La détection de changements nécessite, naturellement, l'acquisition régulière d'informations. Cependant, la fréquence temporelle d'acquisition requise est très variable selon la vitesse d'évolution des phénomènes étudiés. Ainsi, l'analyse des changements globaux se fait plutôt à l'échelle d'années ou de dizaines d'années alors qu'un suivi journalier est nécessaire pour la détection de phénomènes brutaux (inondations, feux, catastrophes naturelles), ou mensuel pour le suivi de la végétation. De plus, l'évolution temporelle étant l'un des facteurs permettant de discriminer le plus précisément les différents types de végétation, la caractérisation des différents types d'occupation du sol nécessite l'acquisition d'informations avec une haute fréquence temporelle.

Aujourd'hui, deux types de capteurs sont considérés pour le suivi des surfaces continentales : les capteurs à haute résolution spatiale (tels que **SPOT/HRVIR**, un pixel pour  $20\text{m} \times 20\text{m}$ <sup>2</sup> mais avec une fréquence d'acquisition mensuelle, et les capteurs avec une fréquence d'acquisition journalière mais une résolution spatiale moyenne ou basse (par exemple, **MeRIS** avec 1 pixel pour  $300\text{m} \times 300\text{m}$  ou **SPOT/VGT** avec 1 pixel pour  $1\text{km}^2$ ). Ces derniers capteurs sont dédiés au suivi de l'occupation des sols et, en particulier, de la végétation. L'exploitation d'informations acquises à différentes résolutions spatiales constitue alors l'un des points clefs pour la classification et la

---

<sup>2</sup>surface acquise en visée nadir.

détection de changements. Plus précisément, il s'agit d'exploiter la richesse de l'information spectrale à haute fréquence temporelle, pour caractériser la végétation et détecter les changements, et la finesse de la résolution spatiale pour extraire l'information structurelle de la scène.

Dans la littérature, de nombreuses méthodes de détection de changements ont été proposées dans le cadre de l'analyse et du suivi des surfaces continentales. Les méthodes qui exploitent des données grand champ, basse résolution spatiale, sont souvent issues de méthodes initialement développées pour l'analyse d'images haute résolution. D'autres reposent sur la comparaison des évolutions temporelles d'une mesure bien choisie (*cf.* analyse de la trajectoire temporelle ou analyse du vecteur de changements 2.3.1). En dehors des limites spatiales de ce type de méthodes, dues à la résolution des données utilisées, la détection par comparaison des évolutions temporelles est généralement, malgré un interétalonnage des images, très sensible aux variations des conditions d'acquisition des données (illumination, date, etc).

Par ailleurs, le problème de la classification a été très exploré dans la littérature mais très peu de travaux ont proposé des solutions pour la classification à une échelle pixellique. De plus, les travaux qui traitent ce problème considèrent généralement de faibles rapports de résolution (2 ou 4) alors que les exigences définies par les applications que nous visons ainsi que les contraintes liées aux données nécessitent de traiter des rapports de résolution de l'ordre de 15 à 50.

## Plan du document

Ce document est organisé en quatre parties.

Dans la partie I, nous présentons l'ensemble des éléments liés au contexte de l'étude et nécessaires pour la définition de la problématique. Nous commençons par présenter les capteurs optiques en activité et leurs principales propriétés avant de préciser certains aspects clefs pour l'étude de la végétation. Ce chapitre montre que l'exploitation de données de différentes résolutions est cruciale pour bénéficier à la fois d'informations discriminantes et d'une bonne localisation spatiale. Dans le chapitre 2, nous effectuons un bref état de l'art sur les méthodes de détection de changements couramment utilisées dans le cadre de diverses applications. Les méthodes dédiées à la détection de changements à partir de séquences d'images basse résolution sont encore peu nombreuses et ne permettent pas, en général, de détecter les changements qui touchent une surface géographique de taille inférieure à celle représentée dans un pixel basse résolution. Afin de permettre ce type de détection, nous effectuons (dans le chapitre 3) une rapide revue des différentes méthodes rencontrées dans la littérature pour une analyse sous-pixellique des images de télédétection. De manière générale, les méthodes développées pour un traitement à une échelle sous-pixellique concernent essentiellement l'estimation de paramètres ou la classification. Cette étude préliminaire nous conduit à poser, dans le chapitre 4, la problématique que nous adoptons pour le suivi de la surface continentale à partir de séquences d'images basse résolution et, en particulier, pour la mise à jour de cartographies.

La deuxième partie de cette thèse est consacrée à la présentation d'une méthode de

classification des surfaces continentales à une échelle sous-pixellique. Le modèle linéaire de mélange, permettant de décomposer la mesure physique acquise en un pixel en fonction des différents types d'occupation du sol représentés au sein de ce pixel, et le critère du maximum *a posteriori* sont à la base de notre approche. Ils conduisent à la définition d'une énergie dont la minimisation permet d'obtenir une classification haute résolution. Une analyse théorique permet de cerner les limites de ce modèle et de mettre en évidence l'information pertinente pour la classification. Elle soulève notamment le problème du choix des dates d'acquisition les plus discriminantes pour la classification. De plus, elle permet de préciser dans quelle mesure ce modèle permet aussi bien de séparer les classes à partir de données basse résolution qu'à partir de données haute résolution. La recherche de l'optimum de cette fonction d'énergie est mise en œuvre par un algorithme de recuit simulé, grâce auquel une classification haute résolution peut être réalisée de manière non-supervisée à partir d'une segmentation haute résolution et d'une séquence d'images basse résolution. Une analyse empirique des performances montre, en particulier, la robustesse de l'approche tant que le rapport de résolution considéré est inférieur à  $20 \times 20$  alors que les méthodes antérieures sont généralement limitées à des rapports de résolution de l'ordre de  $4 \times 4$ .

Dans la partie III, nous considérons le problème de la détection de changements. Pour contourner les problèmes d'interétalonnage couramment rencontrés dans la littérature, nous envisageons le problème de la détection de changements sous la forme d'une comparaison entre une séquence d'images basse résolution à une classification haute résolution. Pour cela, nous nous plaçons dans le cadre probabiliste de la détection *a contrario* récemment introduite en analyse d'images. La méthode que nous proposons ne repose pas sur une modélisation *a priori* des données mais, au contraire, sur le rejet d'un modèle non structuré, dit modèle *a contrario*, par l'observation de données structurées. Ce modèle permet de définir un critère de cohérence entre une classification et une séquence d'images, grâce auquel les changements sont détectés en tant que zones de l'image qui ne contribuent pas au rejet de ce modèle structuré. L'analyse théorique du modèle *a contrario* permet de révéler des propriétés de consistance asymptotiques (lorsque la taille de l'image tend vers l'infini) et non-asymptotiques (par exemple, la détection d'un domaine de taille quelconque dès que le contraste de l'image est suffisamment fort). De plus, nous montrons que ce critère de cohérence, couplé à un algorithme de type RANSAC, conduit à une méthode particulièrement robuste vis à vis de la proportion de pixels de changements présents dans la séquence d'images, facteur généralement limitant pour les méthodes alternatives existantes. Par ailleurs, nous explorons les limites empiriques du domaine d'application de la méthode à la fois en terme de nombre de pixels de changement (la méthode est d'autant plus robuste que ce nombre est petit) et de proportion minimale du changement dans un pixel basse résolution (la méthode est d'autant plus performante que cette proportion est élevée). Par exemple, pour un niveau de contraste standard et dans un contexte monotemporel, les changements impactant plus de 25% d'un pixel basse résolution sont bien détectés tant que moins de 65% des pixels de l'image sont touchés.

La partie IV montre l'apport des méthodes proposées dans le cadre d'un exemple d'application : l'analyse d'une zone agricole de la Plaine du Danube. Dans ce contexte, nous présentons les résultats obtenus de manière indépendante, respectivement pour des applications de carto-

graphie et de détection de changements. D'autre part, nous illustrons au travers de scénarii les collaborations possibles entre détection de changement et classification, telles que la mise à jour et la validation de classifications : la détection de changement peut valider une classification par absence de changement détectés, et la mise à jour de la classification permet de quantifier (en terme de taux d'occupation à l'intérieur des pixels) et qualifier (ou interpréter en terme de type d'occupation du sol) les changements détectés de façon binaire sur les pixels basse résolution.

Première partie .

Contexte de l'étude et objectifs



# 1. Acquisitions satellitaires et observation de la surface terrestre

L'imagerie satellitaire constitue une réserve d'informations variées exploitables dans le cadre de nombreuses applications. Dans ce chapitre, nous présentons les caractéristiques essentielles des systèmes imageurs et, plus précisément, des principaux satellites d'observation de la Terre ainsi que certains aspects spécifiques à l'observation des terres émergées.

## 1.1. Introduction

L'information acquise par des capteurs embarqués sur satellite ou aéroportée correspond à la mesure des signaux électromagnétiques émis ou réfléchis par la surface terrestre. Une image satellite peut être vue comme une application qui associe une mesure, appelée intensité radiométrique, à tout point du domaine spatial observé. Les problématiques étudiées au cours de ce travail de thèse ne sont relatives qu'à que des images acquises par des capteurs optiques, *i.e.* dont la mesure est directement liée à la quantité de lumière solaire réfléchie par la surface ou issue de traitements à partir de telles images. Les intensités radiométriques mesurées dépendent de la surface observée et des caractéristiques de l'instrument d'acquisition.

Une image acquise par un capteur optique mesure, en chaque pixel, la quantité de rayonnement électromagnétique reçue. Ce rayonnement résulte de l'émission de différents rayonnements naturels :

- le rayonnement émis par le Soleil, partiellement réfléchi par la surface terrestre et éventuellement absorbé en partie par l'atmosphère ;
- le rayonnement émis par la surface terrestre et absorbé en partie par l'atmosphère.

Le signal électromagnétique, reçu sous la forme d'un signal analogique par le capteur, est ensuite discrétisé en un signal numérique par un processus d'intégration spatiale et spectrale. Selon les applications visées, les bandes spectrales (nombre et largeur de bandes) sont définies à partir des propriétés d'absorption et de rayonnement des cibles recherchées. La définition des longueurs d'ondes d'observation et le processus d'intégration spectrale s'en déduisent. De même, l'intégration spatiale définit la résolution spatiale des données.

Dans la section 1.2, nous présentons les principales caractéristiques des systèmes imageurs en observation de la Terre de manière à préciser la problématique de l'analyse de la surface conti-



mentale en télédétection. Différents indices introduits dans la littérature pour faciliter la distinction des différents types de végétation sont aussi présentés.

## 1.2. Les capteurs

Les mesures issues d'un capteur satellitaire résultent d'un processus de discrétisation de l'espace observé, elles peuvent être représentées sous la forme d'une grille à deux dimensions (images) ayant des propriétés spécifiques selon l'optique du capteur et la chaîne d'acquisition.

### 1.2.1. Caractéristiques des systèmes imageurs

Pour une application donnée, le choix d'un capteur est généralement déterminé par les résolutions spatiale et spectrale de ses acquisitions, ainsi que par sa fréquence d'acquisition. Ces trois propriétés sont interdépendantes et résultent directement des propriétés optiques du capteur ainsi que de la trajectoire et de l'altitude du satellite sur lequel il est embarqué. Dans cette section, nous rappelons succinctement les principes de base qui permettent de comprendre les contraintes inhérentes aux différents types de capteurs utilisés pour une application donnée.

#### L'orbite et sa fauchée

L'orbite d'un satellite correspond à la trajectoire qu'il effectue autour de la Terre. Cette trajectoire est définie en fonction des objectifs de sa mission et de la capacité des capteurs qu'il transporte. Elle est déterminée par son altitude et son orientation relativement à la Terre.

La surface observée au sol (fauchée) par un capteur est directement liée à son orbite. Il s'agit d'un *couloir* de visée (cf. figure 1.1) dont la largeur varie, selon les capteurs, de 10 km environ (par exemple, **Ikonos**<sup>1</sup>) à 3000 km environ (par exemple, **NOAA/AVHRR 2**). La figure 1.1 offre une vision schématique de l'orbite d'un satellite et de sa fauchée.

Les satellites dits **géostationnaires** sont placés au-dessus de l'équateur, à une altitude (environ 36 000 km) déterminée de manière à avoir une période de rotation égale à celle de la Terre. Ils observent une région fixe de la Terre (cf. figure 1.1). Cette configuration orbitale permet d'acquérir continuellement l'information sur une région donnée. Elle est utilisée, par exemple, pour les applications météorologiques avec les satellites **METEOSAT**<sup>2</sup>. La surface terrestre n'est cependant pas intégralement couverte par des satellites géostationnaires. Par conséquent, cette configuration orbitale n'est pas adaptée aux études liées à l'observation de la Terre dans sa globalité.

L'orbite **quasi-polaire** est définie par une trajectoire d'axe Nord-Sud permettant de couvrir la quasi-totalité de la surface de la Terre avec une périodicité donnée par le couple rotation de

---

<sup>1</sup>Satellite américain pour l'observation de la Terre à très haute résolution (1 à 4m de résolution).

<sup>2</sup>Satellite météorologique européen

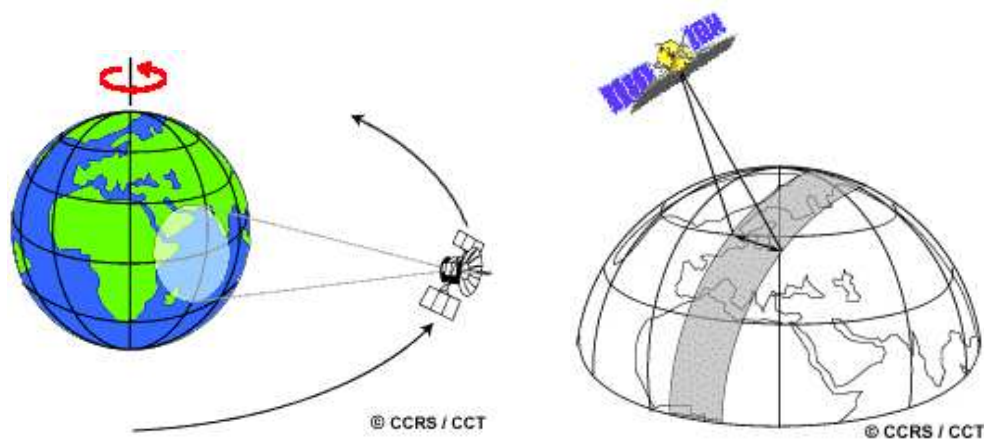


FIG. 1.1.: Représentation schématique de l'orbite d'un satellite géostationnaire (à gauche) et de la fauchée d'un satellite en orbite quasi-polaire (à droite) pour l'observation de la Terre (cf. [CCT, 2001]).

la Terre-orbite du satellite. En général, les satellites placés en orbite quasi-polaire suivent une trajectoire dite héliosynchrone de manière à observer chaque région de la planète à une heure solaire fixée. Les conditions d'illumination solaire sont alors similaires pour chaque acquisition, caractéristique importante pour toutes les applications nécessitant la comparaison d'images successives. Les satellites d'observation de la Terre présentés dans ce chapitre pour l'étude de la surface terrestre sont des plate-formes placées en orbite quasi-polaire, circulaire autour de la planète à une certaine altitude (souvent 800 km) avec un degré d'inclinaison fixé par rapport à l'équateur.

### La résolution spatiale

La résolution spatiale d'un capteur est définie par la distance minimale permettant de séparer deux cibles. En supposant que les pixels des images sont carrés, nous décrivons généralement la résolution spatiale d'un capteur par la largeur de la zone géographique représentée au sein d'un pixel. Le niveau de détail spatial discernable dans une image dépend directement de la résolution spatiale du capteur considéré et des éventuels processus de ré-échantillonnage. Une baisse de résolution spatiale se traduit par une perte d'information quant à l'hétérogénéité spatiale. Les images haute résolution permettent de distinguer les éléments les plus fins, non identifiables par des images basse résolution. La figure 1.2 (a) présente un extrait d'une image **SPOT/HRVIR**, où chaque pixel représente  $20 \times 20 \text{ m}^2$  au sol. Cette résolution permet de bien distinguer les différentes parcelles. Avec une résolution de 300 m (résolution du capteur **MeRIS**), les frontières des parcelles ne sont plus distinguables mais l'allure globale du paysage demeure (cf. figure 1.2 (b)). En revanche, une résolution de 1 km (telle que celle du capteur **SPOT/VGT**) ne laisse réellement apparaître aucun motif (cf. figure 1.2 (c)). Dans ce document, nous appelons *taille du pixel* la

surface au sol observée à l'échelle d'un pixel<sup>3</sup>.

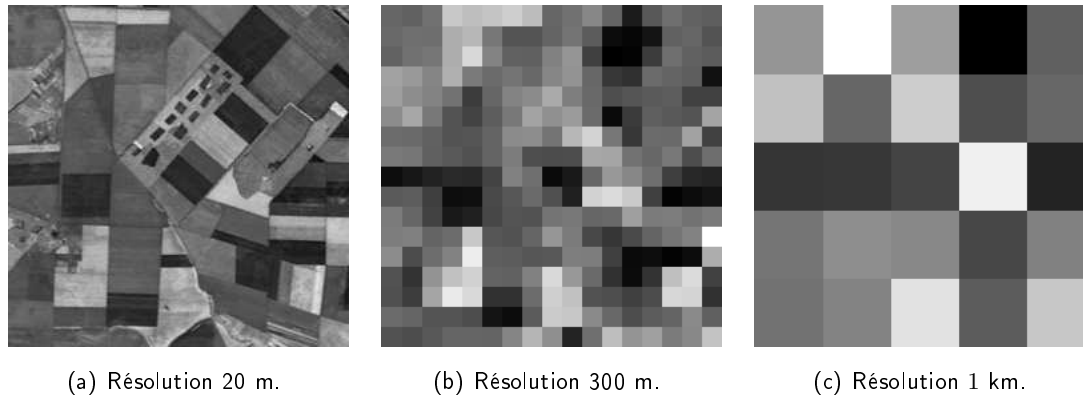


FIG. 1.2.: L'image (a) représente un extrait d'une image de réflectance **SPOT/HRV** de la plaine du Danube acquise en Juin 2001. Les images (b) et (c) représentent des images du même site simulées par moyennage par blocs respectivement de taille  $16 \times 16$  (moyenne résolution, telle que celle de **MeRIS**) et de taille  $50 \times 50$  (basse résolution, telle que celle de **SPOT/VGT**).

### La résolution spectrale

La résolution spectrale est définie par la largeur des bandes spectrales d'un capteur. Elle est d'autant plus fine que les bandes spectrales des différents canaux sont étroites. La résolution spectrale d'un capteur est déterminée précisément pour répondre aux nécessités des applications visées. De plus, si les larges bandes sont suffisantes pour les études d'ordre global, telles que la distinction des grandes classes d'occupation du sol, des bandes plus fines et ciblées sont souvent nécessaires pour les applications plus spécifiques. Par exemple, les applications de types météorologiques nécessitent des informations principalement dans l'infrarouge et l'infrarouge thermique alors que l'analyse de la surface terrestre se fait plutôt à partir de mesures visibles et proche infrarouge. De plus, les applications nécessitant une information radiométrique plus fine (par exemple, pour distinguer des espèces végétales) exploitent généralement des données hyperspectrales, *i.e.* elles utilisent des bandes spectrales plus fines que les applications génériques, mêmes si ces bandes sont toujours situées dans les canaux visibles et infrarouge.

### La résolution temporelle

La résolution temporelle est définie par la période de révolution d'un satellite. En effet, la fauchée des capteurs embarqués permet d'observer la surface terrestre avec une fréquence

<sup>3</sup>La taille d'un pixel varie selon l'angle de visée du satellite mais la connaissance de l'angle de visée et des coordonnées géographiques à l'instant d'acquisition permet une correction géométrique de ces variations.

temporelle de mensuelle (26 jours) pour **SPOT/HRV** à journalière pour **SPOT/VGT**. Même si une fréquence d'acquisition journalière ne semble pas toujours nécessaire, la présence éventuelle de nuages recouvrant une partie de la zone d'observation le jour de l'acquisition est à prendre en compte. Dans la section 1.3, nous verrons l'importance de l'évolution temporelle pour l'analyse et le suivi de la surface terrestre. En effet, si la fréquence temporelle est essentielle pour le suivi de la surface, elle est aussi déterminante pour la distinction des différentes espèces végétales.

Selon l'application visée, les informations requises ne peuvent pas toujours être fournies par un même capteur. En particulier, aucun satellite n'est capable d'apporter à la fois haute résolution spatiale, temporelle et spectrale sur l'ensemble de la surface terrestre. Si les capteurs embarqués sur des satellites géostationnaires peuvent fournir des données précises à la fois spatialement et spectralement, leur position fixée au-dessus de l'équateur ne leur permet pas d'acquérir des informations concernant les pôles.

Dans la section suivante, nous présentons les caractéristiques des principaux satellites et capteurs optiques dédiés à l'observation de la Terre, ainsi que certains aspects essentiels pour l'analyse de la végétation.

### 1.2.2. Les satellites d'observation de la Terre

Nombreux, les satellites d'observation de la Terre sont créés pour remplir des missions parfois très spécifiques. Dans cette section, nous présentons uniquement les satellites et capteurs optiques utilisés dans le cadre de cette étude, ou susceptibles de l'être. Certains capteurs radars sont parfois utilisés pour l'analyse de la végétation mais, en général, l'imagerie radar est plutôt utilisée pour les applications s'intéressant à l'humidité ou à la rugosité. Toutefois, le fait que les images radar (telles que RSO<sup>4</sup>) puissent apporter de l'information à tout moment, quel que soit le niveau d'éclairement et les conditions atmosphériques, offre un complément intéressant à l'imagerie optique, typiquement lorsque les nuages sont abondants.

#### Les systèmes d'observation de la Terre (SPOT)

Le système d'observation de la Terre **SPOT** a été conçu par le Centre National d'Etudes Spatiales pour distinguer des détails de 10 à 20 mètres sur des paysages terrestres avec un suivi régulier. Le premier satellite de la famille **SPOT** a été lancé en 1986. Aujourd'hui, le système **SPOT** comprend trois satellites : **SPOT 2**, **SPOT 4** et **SPOT 5**, lancés respectivement en 1990, 1998 et 2002.

Les satellites du système **SPOT** sont placés en orbite héliosynchrone quasi-polaire à 820 km d'altitude. Ils défilent à une vitesse de 24 000 km/h et sont composés principalement de deux parties : la plate-forme et la charge utile. La plate-forme embarque la charge utile et assure un certain nombre de fonctions liées au fonctionnement du satellite (génère et stocke l'énergie, contrôle

---

<sup>4</sup>Radar à Synthèse d'Ouverture

l'altitude du satellite, etc). La charge utile embarquée est composée de un à quatre instruments de prise de vues, selon les satellites, ayant des propriétés particulières adaptées aux applications visées (cartographie, suivi de la déforestation, agriculture, défense, etc). Les trois premiers satellites de la série **SPOT** contiennent un capteur haute résolution visible **HRV** composé de trois bandes spectrales :

- la bande XS1 couvre les longueurs d'onde comprises entre 0,50 et 0,59  $\mu\text{m}$  (bande d'absorption du vert),
- la bande XS2 couvre les longueurs d'onde situées entre 0,61 et 0,68  $\mu\text{m}$  (bande d'absorption du rouge) et
- la bande XS3 couvrant les longueurs d'onde situées entre 0,78 et 0,89  $\mu\text{m}$  (bande proche infrarouge).

Les satellites **SPOT 4** et **SPOT 5** contiennent une bande supplémentaire, XS4, couvrant les longueurs d'ondes comprises entre 1,58 et 1,75  $\mu\text{m}$  (moyen infrarouge) pour répondre aux besoins spécifiques à l'observation de la végétation. En effet, nous verrons (section 1.3) que le moyen infrarouge permet une analyse fine de l'état hydrique de la végétation. Ces fréquences spectrales ont été choisies pour un maximum d'efficacité dans l'étude et l'analyse des surfaces continentales. Les capteurs haute résolution de ces deux derniers satellites, composés des bandes XS1 à XS4 sont appelés **HRVIR**. Dans le tableau 1.1 sont récapitulées les caractéristiques principales des satellites et capteurs de la famille **SPOT**.

Trois modes de prise de vue sont disponibles pour les instruments **HRV** et **HRVIR** : panchromatique, multispectral ou bien une combinaison des deux modes précédents. Le mode panchromatique correspond à l'utilisation de la bande XS2 uniquement. Il permet d'observer la Terre avec 10 m de résolution (pour les satellites **SPOT 2,3,4**). Le mode multispectral correspond à l'utilisation des quatre bandes spectrales avec une résolution de 20 m. Cette résolution spatiale est utilisée notamment pour décrire l'hétérogénéité spatiale de la couverture végétale.

Lancé le 4 Mai 2002, le satellite **SPOT 5** est le plus récent de la famille **SPOT**. Il apporte une nette amélioration de la résolution spatiale des capteurs **HRVIR** comparativement à ses prédécesseurs même si, par ailleurs, il possède les mêmes caractéristiques. En effet, **SPOT 5** permet d'acquérir des images dans les canaux XS1, XS2 et XS3 avec une résolution spatiale de 10 m alors que **SPOT 4** est limité à une résolution de 20 m. De plus, en mode panchromatique, sa résolution spatiale atteint 5 m, voire 2.5 m en effectuant un traitement *a posteriori* particulier. Un tel niveau de précision ouvre de nouvelles perspectives, notamment pour les applications liées à l'étude du tissu urbain et à la reconnaissance d'objets (types de bâtiments, véhicules, etc).

Avec une fauchée de 60 km, les instruments **HRV** ou **HRVIR** de ces trois satellites ne peuvent observer qu'une surface limitée à un instant donné. Une même surface géographique peut alors être observée en visée nadir<sup>5</sup> tous les 26 jours. Cependant, la présence éventuelle de nuages entre la surface terrestre et le satellite ne permet en pratique qu'un nombre très limité

---

<sup>5</sup>Le nadir désigne la direction perpendiculaire au plan de l'horizon et diamétralement opposée à celle du zénith (i.e. verticale vers le bas).

Satellites	Spot 1,2,3	Spot4	Spot5
Dates de lancement	1986,1990,1993	1998	2002
Altitude	822 km	822 km	830 km
Capteurs embarqués	HRV : V,R,IR	HRVIR : R,V,IR,MIR VGT : B0,R,IR,MIR	HRVIR : R,V,IR,MIR VGT : B0,R,IR,MIR HRG : R,V,IR,MIR,PAN HRS : PAN
Bandes spectrales	V [0, 50 – 0.59 $\mu m$ ] R [0, 61 – 0.68 $\mu m$ ] IR [0, 79 – 0.89 $\mu m$ ] PAN [0, 51 – 0.73 $\mu m$ ]	V [0, 50 – 0.59 $\mu m$ ] R [0, 61 – 0.68 $\mu m$ ] IR [0, 79 – 0.89 $\mu m$ ] PAN[0, 61 – 0.68 $\mu m$ ] MIR [1.58 – 0.1.75 $\mu m$ ] B0 [0, 43 – 0.47 $\mu m$ ]	V [0, 50 – 0.59 $\mu m$ ] R [0, 61 – 0.68 $\mu m$ ] IR [0, 79 – 0.89 $\mu m$ ] PAN[0, 51 – 0.73 $\mu m$ ] MIR [1.58 – 0.1.75 $\mu m$ ] B0 [0, 43 – 0.47 $\mu m$ ]
Résolution spatiale	PAN : 10m HRV : 20m	PAN : 10m HRVIR : 20m  VGT : 1km	PAN : 5m ou 2.5m V,R,IR : 10m MIR : 20m VGT : 1km
Fréquence de passage	26 jours	HRVIR : 26 jours VGT : 1 jour	HRVIR : 26 jours VGT : 1 jour

TAB. 1.1.: Caractéristiques des satellites **SPOT**.

d'observations par an. Par ailleurs, en couvrant une surface au sol d'environ  $60 \times 60 \text{ km}^2$  par image, ces capteurs sont plus appropriés pour des études locales que nationales ou globales.

L'instrument Végétation (**VGT**), embarqué sur les satellites **SPOT 4** et **SPOT 5**, est un capteur *grand champ* complémentaire à **HRVIR**. En effet, ce capteur est programmé pour l'analyse de la végétation à une échelle globale. Il bénéficie donc d'une fauchée de 2250 km au sol au détriment de la résolution spatiale qui est alors de  $1 \text{ km}^2$ . Cette configuration permet une fréquence temporelle d'acquisition journalière, propriété essentielle pour le suivi de la surface terrestre, bien qu'elle soit associée à une résolution spatiale grossière (1 km). Par ailleurs, la résolution spectrale des capteurs **VGT** est identique à celle des capteurs **HRVIR**, à l'exception de la première bande (bande B0 à la place de XS1, cf. tableau 1.1). Remarquons que cette propriété présente un avantage important pour l'utilisation complémentaire de ces deux capteurs, d'autant plus qu'ils bénéficient des mêmes conditions d'acquisition (éclairage, heure, etc). L'interétalonnage entre les données provenant de ces deux capteurs est alors très simple.

Les satellites de la famille **SPOT** apportent donc simultanément des mesures à haute résolution spatiale (**HRVIR**) et à haute fréquence temporelle (**VGT**). Les propriétés spectrales et temporelles du capteur **VGT** en font un instrument particulièrement adapté pour la surveillance de la surface terrestre et l'analyse de la végétation, bien qu'il ne permette pas la localisation précise

des éléments géographiques observés.

### Le capteur superspectral MERIS

Lancé le 1er Mars 2002 et embarqué à bord du satellite **Envisat**, le spectromètre imageur à résolution moyenne (**MeRIS**) a été développé par l'Agence Spatiale Européenne. Le satellite **Envisat** fait partie d'un grand programme européen conçu pour suivre l'évolution de la planète et l'impact des activités humaines : il vise notamment à contrôler le respect des engagements des signataires du Protocole de Kyoto sur la réduction des gaz à effet de serre. Différents capteurs à bord du satellite apportent des informations précises et complémentaires, notamment dans les domaines biophysiques s'intéressant à la surface continentale, à la vapeur d'eau et à la composition chimique des eaux. L'objectif initial de **MeRIS** est l'observation des surfaces marines et côtières et, en particulier, la mesure de la concentration en phytoplancton ou le contrôle de la pollution marine.

Satellite placé (à une altitude moyenne de 800 km) en orbite héliosynchrone quasi-polaire, **Envisat** a un cycle de 35 jours. Le spectromètre **MeRIS** mesure le rayonnement solaire réfléchi par la Terre (et éventuellement les nuages) dans 15 bandes spectrales situées dans le visible et l'infrarouge, dont les centres et la largeur sont programmables. Il permet, avec une résolution spatiale de 300 mètres au sol, une fauchée de 650 km. Avec une résolution spatiale plus grossière (1.2 km), il offre aussi la possibilité d'acquérir des images représentant une surface au sol de  $1150 \times 1150 \text{ km}^2$ . Une même scène peut alors être revisitée tous les trois jours, quel que soit sa situation géographique.

Par la finesse de son spectre, l'instrument **MeRIS** permet de nombreuses applications liées non seulement à l'analyse des surfaces maritimes et côtières (activité biophysique, qualité de l'eau, etc) mais aussi à l'analyse de la végétation terrestre. En effet, sa haute répétitivité temporelle associée à la finesse de son spectre en font un outil particulièrement adapté pour le suivi de la végétation [Justice et Townshend, 2002]. Ces propriétés constituent un avantage important par rapport aux autres capteurs optiques.

D'autres satellites sont dédiés à l'analyse de la surface terrestre. En particulier, le satellite américain **Landsat 7** fournit des données multispectrales sur la surface terrestre avec une résolution spatiale de 30m et un cycle de 16 jours. Le capteur américain **NOAA/AVHRR**, initialement dédié aux études océanographiques et atmosphériques (National Oceanographic and Atmospheric Administration), peut aussi être utilisé pour l'analyse de la surface continentale ou de la végétation à une échelle globale. En effet, ses caractéristiques spectrales et temporelles en font un outil adapté aussi pour l'analyse de la végétation aux échelles globale et régionale, comme le capteur **SPOT/VGT**, notamment pour l'étude des changements de saison : il permet d'acquérir deux fois par jour des images de la même scène avec une résolution de 1.1 km.

## 1.3. Spécificités liées à la végétation

Depuis plusieurs décennies, la télédétection est utilisée pour la classification et le suivi de la végétation ainsi que pour des analyses plus spécifiques dans le cadre, notamment, de l'agriculture de précision ou de la surveillance des forêts. Cependant, les enjeux sont différents selon la taille de la zone étudiée : si l'agriculture de précision consiste souvent à analyser les hétérogénéités intra-parcellaires dans l'objectif d'adapter les pratiques agricoles aux besoins sous-parcellaires et d'optimiser ainsi la production, les études à l'échelle nationale, européenne ou globale nécessitent un niveau moins fin d'analyse de la végétation mais sur une surface plus vaste. L'analyse est alors liée simplement à la distinction des principaux types d'occupation du sol plutôt qu'aux variations au sein d'un type de culture ou d'une parcelle.

Dans cette section, nous présentons les principales caractéristiques spectrales de la végétation avant d'introduire quelques indices de végétation très utilisés dans la littérature pour discriminer les différentes espèces végétales et évoquons l'importance de l'aspect temporel pour ce type d'applications.

### 1.3.1. Les propriétés spectrales

Avant de préciser les bandes spectrales les plus adaptées pour l'analyse de la végétation, rappelons les bases des interactions entre la surface terrestre et la canopée<sup>6</sup>. En supposant qu'une couche de végétation recouvre le sol, les rayons lumineux qui atteignent la canopée sont partiellement réfléchis par la végétation mais aussi partiellement transmis et diffusés par la végétation en direction du sol, où ils sont réfléchis. Les rayons réfléchis par le sol sont, à leur tour, partiellement transmis par la végétation et partiellement réfléchis par la végétation vers le sol. La figure 1.3 représente les différents cas de transmission et de réflexion par le sol et la végétation d'un rayon solaire.

De nombreux modèles de transfert radiatif dans l'écosystème terrestre proposent une formulation des relations entre les mesures de télédétection et les caractéristiques biophysiques de la végétation. Ils permettent de relier un ensemble de paramètres biophysiques aux mesures multispectrales acquises par télédétection. Notamment, les modèles SAIL (*cf.* [Verhoef, 1984]) ou Adding (*cf.* [Cooper et al., 1982]) sont tous les deux fondés sur la théorie du transfert radiatif de [Chandrasekhar, 1950] qui fournit une bonne approximation des processus de diffusion quand les diffuseurs sont grands devant la longueur d'onde. L'inversion de ce type de modèles permet d'accéder à des paramètres de végétation à partir des mesures de télédétection.

Les propriétés optiques des sols sont liées essentiellement à leur composition minérale. En effet, le spectre de réflectance d'un sol correspond à la superposition des spectres de chaque composante minérale du sol. Comme les minéraux, les sols ont une réflectance croissante du visible au proche infrarouge, avec des bandes d'absorption autour de  $1.4\mu\text{m}$  et de  $1.9\mu\text{m}$ . L'humidité du

---

<sup>6</sup>Etage supérieur de la végétation, en contact avec l'atmosphère.



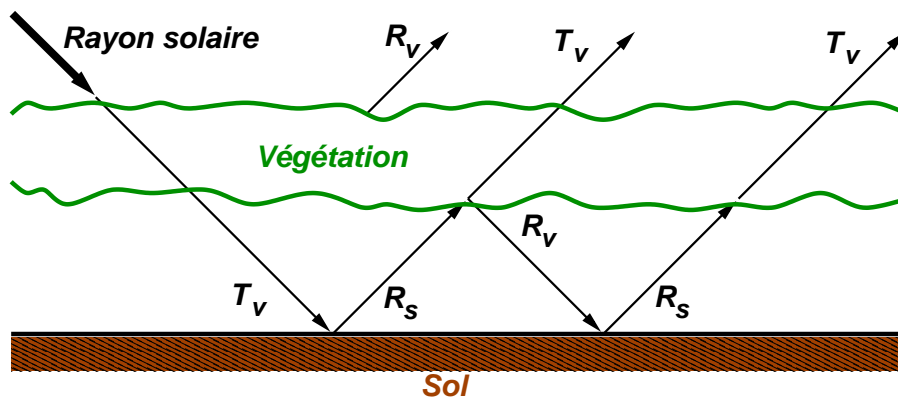


FIG. 1.3.: Interactions entre la couche de végétation et la surface terrestre. Les rayons lumineux notés  $R_v$ ,  $R_s$  et  $T_v$  désignent, respectivement, les rayons réfléchis par la végétation, par le sol et transmis par la végétation.

sol constitue aussi un facteur affectant les propriétés optiques des sols, ainsi que le contenu en matière organique. Ce dernier a un effet sur les propriétés optiques dans les bandes inférieures à  $1.8\mu\text{m}$ .

Pour l'étude de la végétation, il est essentiel de pouvoir caractériser les principaux aspects de la canopée :

- l'absorption par la chlorophylle,
- le contenu en eau, et
- les propriétés structurelles.

La figure 1.4 présente de manière schématique les réponses spectrales typiques de la végétation et du sol nu. Elle indique également les bandes spectrales rouge, vert, proche infrarouge et moyen infrarouge. Les bandes spectrales d'acquisition les plus utilisées sont la bande rouge, centrée sur le pic d'absorption de la chlorophylle ( $0.665\mu\text{m}$ ) et la bande proche infrarouge, correspondant à la réflectance maximum de la végétation et liée aux propriétés structurelles de la canopée et du pourcentage de sol nu couvert de végétation. L'utilisation supplémentaire d'une bande moyen infrarouge centrée autour de  $1.65\mu\text{m}$  permet de prendre en compte le contenu en eau des composants de la canopée. La bande spectrale correspondant au moyen infrarouge ( $1.58 - 1.75\mu\text{m}$  pour **SPOT**) est particulièrement adaptée pour caractériser la couverture végétale. Les feuilles et le sol ont des propriétés optiques particulières qui se traduisent par de très faibles valeurs dans le visible. En effet, les pigments présents dans les feuilles absorbent fortement les rayons provenant de la bande rouge. Dans le proche infrarouge, le phénomène d'absorption est absent et les valeurs observées sont alors saturées. Le moyen infrarouge, en revanche, conduit à l'observation de mesures de réflectance intermédiaires dont la variabilité permet une meilleure discrimination des différents types de végétation. Il est très sensible aux variations structurelles du couvert végétal. De plus, l'atmosphère apparaît comme très transparente dans cette bande. Les perturbations associées à la variabilité des conditions atmosphériques sont alors moindres.

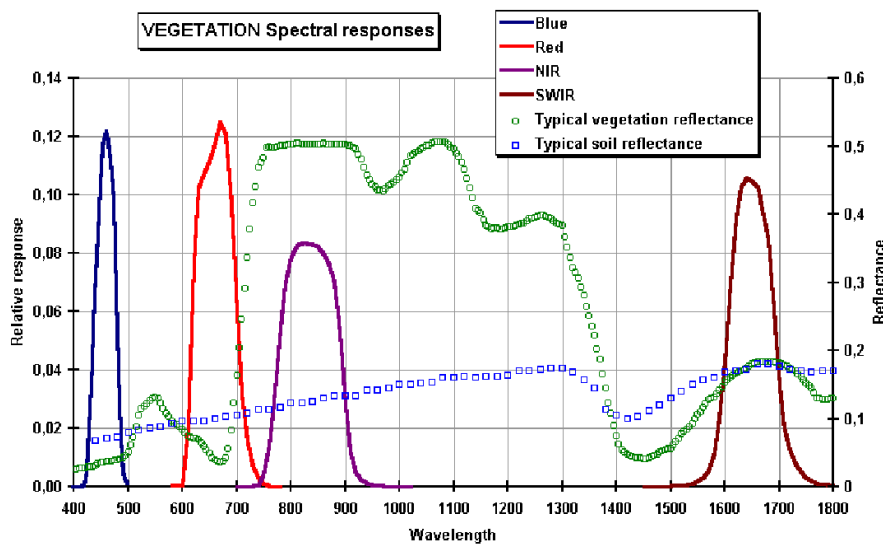


FIG. 1.4.: Les réponses spectrales de la végétation (cf. [CCT, 2001]) : mesure de réflectance en fonction de la longueur d'ondes (en nanomètres).

Notons que la résolution spectrale requise est étroitement liée à l'application et, en particulier, à la précision souhaitée pour le suivi de la végétation : les études à l'échelle de la parcelle agricole (par exemple pour l'agriculture de précision, qui nécessite de discriminer des paramètres biophysiques de la végétation au sein d'une même culture) exigent, en général, une fine résolution spectrale alors que l'objectif d'une analyse de la surface à l'échelle nationale ou globale nécessite une résolution spectrale moindre. En effet, il suffit alors généralement de discriminer les grands types de végétation.

### 1.3.2. Les indices de végétation

Les réflectances mesurées directement par un capteur dans différentes bandes spectrales offrent des informations redondantes pour l'analyse de la végétation. Les indices de végétation proposent des combinaisons (linéaires ou non) des réflectances mesurées dans plusieurs bandes spectrales, en particulier dans les bandes rouge (**R**) et proche infrarouge (**PIR**), ciblées pour la caractérisation du couvert végétal. Ils permettent de réduire les effets des conditions d'acquisition (éclairage solaire, propriétés optiques du sol, etc) par rapport à la valeur radiométrique observée. La littérature propose de nombreux indices spécifiques et sensibles aux caractéristiques du sol et, plus ou moins, aux conditions atmosphériques.

Un indice de végétation est construit dans l'objectif de réduire les mesures multispectrales à une seule valeur informative pour la prédiction et l'évaluation de caractéristiques de végétation. Il existe différentes formules d'indices de végétation, généralement empiriques. En général, ces indices correspondent à la combinaison des mesures acquises dans le canal rouge (**R**), ondes absorbées par la chlorophylle pour la photosynthèse, et dans le proche infrarouge (**PIR**), ondes réfléchies par

la surface des feuilles.

L'indice de végétation le plus ancien est le RVI (Ratio Vegetation Index), introduit par [Pearson et Miller, 1972]. Il est défini par

$$RVI = \frac{PIR}{R},$$

où **PIR** représente la réflectance dans la bande proche infrarouge et **R** correspond à la réflectance dans la bande rouge (visible). Lorsque la surface de végétation verte augmente, le dénominateur diminue et le numérateur augmente. Cet indice varie globalement de 1 pour les sols nus à plus de 20 pour la végétation dense (*cf.* [Knippling, 1970]). Cet indice permet d'accentuer le contraste entre la végétation et le sol, il est peu affecté par les conditions d'illumination. En revanche, il est sensible aux effets atmosphériques et très peu sensible aux faibles taux de couverture. En général, il semble donc plutôt adapté à la végétation dense.

D'autres indices ont été proposés dans la littérature, fondés sur la différence entre les réflectances obtenues dans les bandes **PIR** et **R** afin de compenser les différents effets de la lumière entrante. Ils sont souvent normalisés de manière à fournir un indice compris entre 0 et 1.

L'indice le plus couramment utilisé pour l'imagerie satellite est le NDVI (Normalized Difference Vegetation Index), lié à la fraction du rayonnement actif pour la photosynthèse et donc à la productivité et à la biomasse de la végétation. Cet indice a été introduit en 1973 par [Rouse et al., 1973] pour identifier rapidement et simplement les surfaces végétales. Depuis, divers travaux justifient expérimentalement son utilisation dans différents cas d'application. En particulier, [Vignolles, 1996] étudie son efficacité pour les applications agricoles. Il est défini par

$$NDVI = \frac{PIR - R}{PIR + R}.$$

Le NDVI vaut, typiquement, 0.1 pour les sols nus et 0.9 pour la végétation dense. Il est considéré comme plus sensible aux faibles niveaux de végétation qu'à la végétation dense, contrairement au RVI qui est plus sensible aux variations de la végétation lorsqu'elle est dense. En revanche, le NDVI est sensible, comme le RVI, aux conditions d'illumination pour les faibles densités de couvert, aux effets atmosphériques et à l'angle de visée. Cependant, il est très utilisé en télédétection et souvent considéré comme pertinent pour caractériser les différents types de végétation, grâce à son évolution temporelle. En effet, l'évolution de ce type d'indice de végétation correspond au développement phénologique des espèces végétales observées.

Ce type d'indice est construit de manière à obtenir de fortes valeurs pour les surfaces de plantes vertes. En effet, la chlorophylle (pigment vert) absorbe le rayonnement incident de la partie visible du spectre alors que la structure des feuilles et la teneur en eau conduisent à de fortes valeurs de réflectance dans la bande **PIR**. De nombreux dérivés ou alternatives au NDVI ont été proposés dans la littérature pour outrepasser ses limites, en particulier en tenant compte de propriétés géologiques des sols pour créer un indice *corrigé* des effets dûs au type de sol. Plus précisément, la *droite des sols* correspond à une représentation du sol dans le plan (**R**,**PIR**). La

Indice	Abréviation	Formule
Ratio vegetation index [Pearson et Miller, 1972]	RVI	$\frac{\text{PIR}}{\text{R}}$
Normalized vegetation index [Rouse et al., 1973]	NDVI	$\frac{\text{PIR}-\text{R}}{\text{PIR}+\text{R}}$
Perpendicular vegetation index [Richardson et Wiegand, 1977]	PVI	$\frac{\text{PIR}-a\text{R}-b}{\sqrt{a^2+1}}$
Soil-adjusted vegetation index [Huete, 1988]	SAVI	$\frac{(\text{PIR}-\text{R})(1+L)}{\text{PIR}+\text{R}+L}, L = 0.5$
Transformed soil-adjusted vegetation index [Baret et al., 1989, Baret et Guyot, 1991]	TSAVI	$\frac{a(\text{PIR}-a\text{R}-b)}{a\text{PIR}+\text{R}+X(1+a^2)}, X = 0.08$
Modified soil-adjusted vegetation index [Qi et al., 1994]	MSAVI	$\frac{2\text{R}+1-\sqrt{(2\text{R}+1)^2-8(\text{R}-\text{PIR})}}{2}$

TAB. 1.2.: Principaux indices de végétation.

végétation est alors d'autant plus dense que sa représentation dans ce plan est éloignée de la *droite des sols*.

Dans cet esprit, [Richardson et Wiegand, 1977] ont proposé le PVI (Perpendicular Vegetation Index), défini par la plus courte distance entre un point du plan  $(\text{R}, \text{PIR})$  à la *droite des sols* (d'équation  $\text{PIR} = a\text{R} + b$ ) de manière à amoindrir les effets optiques du sol. Cependant, le PVI est fortement limité par les effets atmosphériques et la géométrie d'acquisition. Il a ensuite été amélioré par [Huete, 1988] avec le SAVI (Soil-Adjusted Vegetation Index) qui tente de corriger les perturbations imputables à la variabilité des caractéristiques du sol. Le SAVI est défini par

$$\text{SAVI} = \frac{\text{PIR} - \text{R}}{\text{PIR} + \text{R} + L} \times (1 + L),$$

où  $L$  est un terme de correction fixé compris entre 0 et 1 selon le type de sol, mais fixé par défaut à 0.5. D'autres améliorations ont suivi, pour corriger la brillance du sol, notamment avec le TSAVI (*cf.* [Baret et al., 1989]) et le MSAVI (*cf.* [Qi et al., 1994]) qui est moins influencé par les variations spatio-temporelles des sols mais reste sensible aux conditions atmosphériques et d'illumination.

Tous ces indices de végétation, couramment utilisés dans la littérature, présentent des avantages et des inconvénients liés aux conditions expérimentales pour lesquelles ils ont été définis. Le tableau 1.2 fournit un récapitulatif des principaux indices de végétation. Remarquons que le TSAVI est équivalent au NDVI lorsque les paramètres de la *droite des sols* sont  $a = 1$  et  $b = 0$ .

Depuis les années 1990, une nouvelle génération d'indices de végétation est proposée pour estimer directement les variables biophysiques d'intérêt. En particulier, les modèles SAIL et PROSPECT couplés (*cf.* [Jacquemoud et al., 1995]) permettent d'exprimer un ensemble de variables biophysiques en fonction de la mesure multispectrale observée. Nous citons ci-dessous quelques unes de ces variables biophysiques :

- l'indice foliaire LAI (Leaf Area Index) est un indice biophysique correspondant à la surface totale de feuilles vertes par unité de surface au sol ;
- l'angle foliaire moyen MLA (Mean Leaf Angle) est défini par l'angle moyen formé entre les feuilles et l'horizontal. Typiquement, le MLA des cultures à larges feuilles varie entre 20° et 40° alors que celui des céréales varie entre 50° et 60° ;
- la fraction de couvert en visée nadir est une variable biophysique dérivée de l'indice foliaire (LAI) et de l'angle foliaire (MLA). Elle caractérise la proportion de sol couvert de feuillage dans un pixel. Cette variable est, par définition, linéaire par changements d'échelle (rapport de deux surfaces).

[Baret et Guyot, 1991] proposent une équation reliant un indice de végétation au LAI, permettant notamment de convertir le NDVI en LAI. De même, la fraction de couvert peut être calculée en tant que fonction exponentielle du NDVI. Globalement, les profils d'évolution temporelle obtenus à partir du NDVI sont alors proches de ceux obtenus par inversion du modèle biophysique. Cependant, un effet de saturation est observable pour les couverts végétaux denses, effet résultant directement d'une limite intrinsèque au NDVI.

Le NDVI, souvent considéré comme pertinent pour caractériser les différents types de végétation (*cf.* [Fisher, 1994]), est très utilisé par la communauté de télédétection. Cependant, le fait qu'il ne soit pas linéaire n'en fait pas l'indice idéal *a priori* pour l'analyse de la surface à partir d'images de différentes résolutions spatiales. Même si quelques travaux montrent que le NDVI peut être utilisé au-delà de ses limites *théoriques*, *i.e.* comme s'il était linéaire, nous préférons considérer un indice linéaire tel que la fraction de couvert.

### 1.3.3. L'aspect temporel

L'évolution des mesures acquises sur la surface ou celle des indices de végétation au cours du temps constitue une information essentielle pour le suivi de la surface terrestre et, en particulier, pour discriminer les différents types d'occupation du sol. En effet, un type de culture peut être caractérisé à un instant donné par une signature spectrale ou un indice de végétation qui correspond à un stade phénologique<sup>7</sup> particulier. Si certaines cultures peuvent être à peine initiées lors de l'acquisition, d'autres sont déjà sénescentes<sup>8</sup> ou en période de floraison (cultures d'été, d'hiver, etc). Cependant, cette signature spectrale peut être commune, à un instant donné, à différents types de cultures.

En revanche, la signature spectrale de chaque type de végétation a une évolution différente au cours du temps. Même si les limites d'une parcelle agricole peuvent être déterminées à partir d'une seule acquisition, une étude temporelle est généralement nécessaire pour identifier la nature du couvert de cette parcelle.

---

<sup>7</sup>Stade du cycle de vie d'un organisme (animal ou végétal)

<sup>8</sup>La sénescence correspond au processus de dégénérescence cellulaire qui suit la période de maturation ou de vitalité maximale.

L'évolution temporelle d'un indice de végétation est importante dans la mesure où elle révèle l'évolution de la biomasse de la couverture végétale. En effet, la phase ascendante correspond à la croissance de la végétation, la phase stationnaire correspond à la floraison ou maturation des fruits et la phase descendante correspond à la sénescence puis à la récolte. Le cycle phénologique de chaque culture suit un calendrier spécifique. L'évolution temporelle d'un indice de végétation permet donc de différencier les différentes cultures. La figure 1.5 présente l'évolution de la fraction de couvert mesurée en un pixel au cours d'une année agricole (27 dates réparties du 15 Octobre 2000 au 22 Juillet 2001). Les trois graphes correspondent à trois pixels différents de la même séquence.

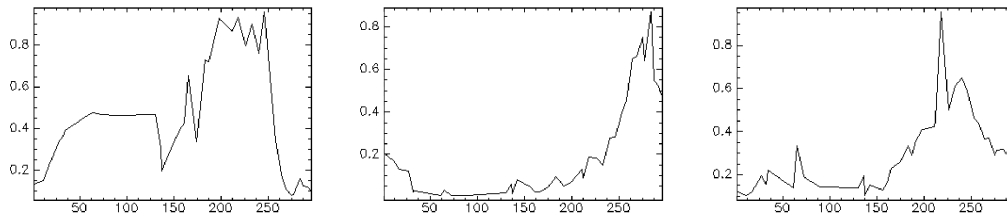


FIG. 1.5.: Profils temporels de fractions de couvert : le temps est représenté en abscisse et la fraction de couvert en ordonnée. Chaque graphe représente l'évolution temporelle de la fraction de couvert observée en un pixel de l'image.

L'évolution temporelle de la signature spectrale, d'un indice de végétation ou d'un paramètre biophysique constitue, par conséquent, une information essentielle pour discriminer les différents types de culture.

## 1.4. Conclusion

Au cours de ce chapitre, les données satellitaires et leurs principales caractéristiques ont été exposées, ainsi que des informations essentielles pour l'étude de la végétation. Pour l'analyse de la végétation, l'idéal serait de disposer à la fois d'images acquises à haute résolution spatiale, pour permettre la localisation des objets d'intérêt, et d'images à haute résolution temporelle et spectrale, pour le contrôle de la surface terrestre et la distinction des différents types de végétation.

La fréquence d'acquisition d'un satellite et sa résolution spatiale sont directement liées à la fauchée du satellite (*cf.* section 1.2.1). Une large fauchée correspond à l'observation d'une grande surface au sol. Les capteurs de ce type (*grand champ*), permettent d'observer une même scène sans nécessairement passer à sa verticale. La fréquence d'acquisition est alors augmentée. En revanche, la visée grand champ est incompatible avec la distinction des détails de la scène. Elle offre une vue globale, particulièrement adaptée aux études liées à l'environnement ou à l'analyse des interactions géosphère-biosphère-atmosphère. Les capteurs grand champ sont donc caractérisés par une haute fréquence temporelle couplée à une faible résolution spatiale. A l'inverse, lorsque la

fauchée est étroite, l'acquisition d'une même scène n'est possible que lorsque le satellite passe à sa verticale. La fréquence temporelle d'acquisition est alors inférieure à celle des données grand champ mais la résolution spatiale est plus fine.

Les caractéristiques inhérentes aux capteurs conduisent donc à un compromis entre finesse de résolution spatiale et finesse de résolution temporelle, elles ne permettent pas d'accéder simultanément à une résolution fine dans ces différents domaines. Les deux types de capteurs **SPOT/HRV** et **SPOT/VGT** sont alors complémentaires. Dans l'objectif de suivre l'évolution de la surface terrestre et, en particulier, de la végétation, nous proposons d'exploiter la complémentarité des capteurs haute résolution spatiale et haute fréquence temporelle, tels que **SPOT/HRV** et **SPOT/VGT**, **MeRIS** ou **NOAA/AVHRR**. L'utilisation conjointe des données provenant des capteurs **HRV** ou **VGT** d'un satellite **SPOT** facilite les prétraitements et permet d'en limiter les erreurs dans la mesure où ces capteurs, embarqués sur un même satellite, bénéficient de paramètres géométriques comparables. Il semble alors raisonnable de supposer qu'une image **VGT** correspond à la moyenne d'une image **HRV** même si les distorsions liées aux angles de visée sont plus disparates pour la basse résolution que pour la haute résolution. Par ailleurs, le capteur **MeRIS** propose une résolution spatiale nettement plus fine que **VGT** pour une fréquence temporelle d'acquisition suffisante pour la classification de la végétation et son suivi.

La plupart des méthodes existantes pour l'analyse et la cartographie de la végétation reposent essentiellement sur l'exploitation de données multispectrales haute résolution. L'information haute résolution spatiale est, en effet, essentielle pour la cartographie dans la mesure où elle permet de localiser les différents objets géographiques (parcelles agricoles, forêts, fleuves, villes, prairies, etc). Cependant, la fréquence des acquisitions est également essentielle pour distinguer les différents types de végétation. De plus, le développement d'outils de suivi à partir de données d'imagerie grand champ apparaît comme un objectif majeur pour une gestion à grande échelle.

Dans ce contexte, notre objectif est de développer des outils automatiques de classification et de détection de changements adaptés au suivi de la surface terrestre et capables d'exploiter conjointement des données provenant de capteurs haute et basse résolution spatiale. En effet, l'exploitation d'informations provenant de différentes résolutions est alors cruciale pour bénéficier à la fois d'une bonne discrimination entre les types de couvert et d'une bonne localisation des objets d'intérêt. Dans le chapitre 2, nous présentons quelques éléments d'état de l'art en détection de changements. La plupart des méthodes rencontrées dans la littérature proposent des solutions pour la détection de changements entre deux dates à partir d'images haute résolution ou entre deux périodes à partir d'images basse résolution. Dans le chapitre 3, nous abordons certains aspects spécifiques à l'utilisation conjointe d'images haute et basse résolution.

## 2. Méthodes de détection de changements

Avant d'envisager le problème de la détection de changements dans le cadre spécifique du suivi de la surface terrestre et, en particulier, des surfaces continentales, ce chapitre est centré sur les différentes méthodes utilisées pour la détection de changements à partir de données images. Le problème de la détection de changements en traitement d'images est motivé par des applications très diverses. Nous proposons ici un état de l'art sommaire des techniques couramment utilisées afin d'en souligner les principaux avantages et inconvénients. Pour finir, nous discutons de la possibilité de les appliquer dans le contexte particulier du suivi de la végétation, en regard des données à notre disposition.

### 2.1. Introduction

#### 2.1.1. Problématiques de la détection de changements

L'essor des technologies liées à l'imagerie satellite se manifeste par une croissance marquée de la production d'images satellites et une mesure de plus en plus ciblée vis à vis des nombreuses problématiques concernant l'observation, la protection ou la surveillance de la planète. De nombreuses problématiques abordées aujourd'hui en télédétection sont liées à la détection de changements, c'est-à-dire à la caractérisation et à la localisation des zones qui ont évolué entre deux instants ou deux périodes données (à partir de deux observations ou séquences d'observations) sur une même scène. Cette évolution peut être naturelle et, par exemple, correspondre à la croissance de la végétation, à une éruption volcanique, à une inondation, à un incendie ou à un glissement de terrain. Elle peut aussi être liée à l'activité humaine de manière directe, avec l'urbanisation, les coupes de forêt, ou la rotation des cultures, ou de manière indirecte, notamment à travers les effets de la pollution. En effet, la pollution a de fortes répercussions sur le climat et la surface terrestre. Les changements induits sont donc de différents types, d'origines et de durées variées. Nous pouvons, à titre d'exemple, distinguer les familles d'applications suivantes :

- le suivi du plan d'occupation des sols, notamment pour des applications en géographie (urbanisation, déforestation, politique agricole),
- la détection des changements saisonniers de végétation, pour des applications en agromonie,
- la gestion des ressources naturelles, notamment le suivi des opérations d'afforestation, de reforestation ou de déforestation (ARD) dans le cadre des protocoles internationaux



(par exemple, le Protocole de Kyoto),

- la cartographie des dommages, qu'ils soient dus aux catastrophes naturelles (inondations, raz de marée, éruptions volcaniques, tremblements de terre, etc), ou d'origine anthropique.

La détection de changements repose essentiellement sur la capacité à mesurer les aspects temporels des phénomènes en utilisant des données multitemporelles bien choisies. L'étude de [Edwards, 1990] révèle qu'une interprétation visuelle d'une photographie aérienne produit presque toujours de meilleurs résultats que les méthodes numériques de détection automatique des changements, avec un degré de précision plus élevé. Cependant, outre leur coût élevé, les résultats d'une interprétation visuelle demeurent subjectifs ce qui incite au développement de techniques de détection automatique, même s'il s'agit d'une tâche difficile [Coppin et al., 2004].

En traitement d'images, la détection des zones de changements dans des paires ou séquences d'images d'une même scène prise à différentes dates est un problème qui concerne de nombreux domaines, tels que l'aide au diagnostic médical [Rey et al., 2002, Bosc et al., 2003], la vidéo-surveillance [Wren et al., 1997, Cao et al., 2005] ou la télédétection [Collins et Woodcock, 1996, Bruzzone et Prieto, 2002, Le Hégarat-Masclé et Seltz, 2004]. Il s'agit d'identifier l'ensemble des pixels qui diffèrent *significativement* entre deux instants ou deux périodes données. Le problème est généralement formulé de la manière suivante : à partir de deux images (ou séquences d'images)  $I_1$  et  $I_2$  acquises à deux instants (ou intervalles de temps)  $t_1$  et  $t_2$  différents, l'objectif est de générer une image représentant les zones de changements/non-changement entre  $I_1$  et  $I_2$ , image couramment appelée carte des changements. Le problème de la détection de changements a donné lieu à de nombreux travaux. Il est souvent décomposé en deux étapes :

1. l'élaboration d'un critère ou d'une mesure permettant de faciliter la discrimination des changements, c'est-à-dire de mettre en exergue les changements survenus entre les deux périodes considérées ;
2. l'identification des changements de manière à obtenir une partition de la scène observée en zones de changements et de non-changement (prise de décision).

La génération de données révélatrices (*cf.* section 1.3.2) des changements apparus entre les deux instants ou périodes d'observation est donc un point clef pour la qualité de la détection. L'utilisation de critères calculés directement sur les images originales n'est, en effet, pas forcément performante. Par exemple, certains types d'images peuvent être très sensibles aux conditions d'acquisition : dans le cas des images optiques, l'information radiométrique est très sensible aux conditions d'illumination et à l'angle de visée.

### 2.1.2. Taxonomie des méthodes de détection de changements

Dans les sections suivantes, nous présentons un état de l'art synthétique sur la détection de changements appliquée à la télédétection. L'objectif n'est pas, ici, de dresser une liste exhaustive des méthodes de détection de changements proposées dans la littérature depuis plusieurs décennies,

mais plutôt de mettre en avant différentes stratégies pour réaliser la détection. Les articles d'état de l'art de [Singh, 1989], [Jensen et Cowen, 1997], [Lu et al., 2004], ou [Coppin et al., 2004] proposent une revue plus complète des approches existantes de détection de changements en télédétection. Notons que les catégories utilisées pour classifier les méthodes de détection diffèrent d'un article à l'autre. En particulier, [Coppin et al., 2004] proposent une taxonomie distinguant le cas bi-dates, *i.e.* la détection de changements réalisée uniquement entre deux images acquises à deux instants différents, du cas multi-dates, couramment appelé *temporal trajectory*, dont l'objectif est de réaliser le suivi d'une scène à partir d'une série d'observations sur une longue durée (évolution des cultures, du tissu urbain, etc)<sup>1</sup>. Cependant, les méthodes d'analyse de la trajectoire temporelle sont très peu nombreuses et correspondent généralement à des adaptations de méthodes bi-dates au cas multi-dates. Par la suite, nous qualifions de bitemporelles les méthodes de détection de changements basées sur la comparaison entre deux dates, et de multitemporelles celles qui reposent sur la comparaison entre plusieurs dates. En général, les méthodes bitemporelles sont plutôt employées pour l'analyse des changements abrupts, à une échelle locale, et exploitent des données à haute résolution spatiale. A l'inverse, les méthodes multitemporelles, plutôt centrées sur l'analyse des évolutions lentes ou des phénomènes à grande échelle, reposent sur l'analyse des profils d'évolution à haute fréquence temporelle dérivés de séries *grand champ* acquises par des capteurs à basse résolution spatiale et haute répétitivité temporelle. Pour l'analyse de l'évolution de la végétation et des attributs de surface à une échelle régionale ou nationale voire globale, des outils de détection capables d'exploiter l'information multitemporelle sont nécessaires. Cependant, la résolution spatiale des capteurs à haute fréquence temporelle est un facteur très limitant pour le type de changements observables. Par ailleurs, [Coppin et al., 2004] identifient, dans le cas bi-dates, neuf grands types de méthodes : la comparaison post-classification, l'analyse conjointe, les opérations de différence et de rapport d'images, les transformations linéaires, l'analyse de vecteurs de changements, la régression, l'analyse multitemporelle de mélange spectral, l'analyse multidimensionnelle de l'espace temporel et les méthodes hybrides plus marginales. Dans un cadre plus général, [Radke et al., 2005] présentent une revue très complète des principales méthodes de détection de changements, aussi bien à partir d'images satellites que d'images beaucoup plus proches dans le temps, issues par exemple de séquences vidéo.

Dans ce chapitre, nous proposons de distinguer, en adoptant un critère de classification davantage méthodologique, deux grandes classes de méthodes de détection de changements :

- les méthodes basées sur les différences observées d'après la mesure du capteur en chaque point de l'image, qualifiées naturellement d'approches ponctuelles ;
- les méthodes basées sur l'extraction et le suivi de régions ou d'objets dans les images, regroupées sous le nom d'approches contextuelles.

Notons, cependant, que nous ferons souvent référence à [Coppin et al., 2004] pour la présentation de méthodes d'intérêt bien que nous n'ayons donc pas adopté sa taxonomie.

---

<sup>1</sup>Notons qu'une approche est également qualifiée de bi-dates lorsqu'elle est basée sur l'utilisation de deux séquences d'images acquises sur deux intervalles de temps donnés.

Avant d'appliquer une quelconque technique de détection de changements à des images de télédétection, un certain nombre de prétraitements sont nécessaires pour rendre les images comparables : les principaux d'entre eux sont présentés brièvement dans la section 2.2.

## 2.2. Prétraitements

La détection de changements est possible seulement si les changements sur le phénomène d'intérêt impliquent des changements de réflectance. De plus, si un capteur est suffisamment sensible pour détecter un changement, l'objectif est généralement de séparer les changements liés au phénomène d'intérêt des autres changements et surtout du bruit ou de la variabilité présent dans les images. Les facteurs qui peuvent rendre difficile cette distinction sont les différences liées aux conditions atmosphériques ou à l'éclairement (angle d'incidence, date d'acquisition, etc.), les différences des géométries d'acquisition et les différences de l'état de la surface (humidité).

Avant d'appliquer une méthode de détection de changements, il paraît donc essentiel de pouvoir assurer, au cours d'une étape de prétraitements, qu'une comparaison est possible entre les deux observations, tant au niveau géographique (les images correspondent bien à la même zone observée) qu'au niveau de la qualité radiométrique des images (les observations ont été faites dans des conditions similaires). En particulier, les images utilisées dans le cadre d'une analyse multitemporelle sont souvent acquises dans des conditions différentes, notamment concernant l'atmosphère, qui impliquent des variations de réflectance. Des changements de radiométrie dus à la présence de nuages *fins* et à leurs ombres arrivent aussi fréquemment, ils sont sans rapport avec les changements qui nous intéressent mais risquent d'être perçus comme tels en l'absence d'un pré-traitement adéquat. Il est donc essentiel de corriger au préalable la radiométrie des effets atmosphériques, mais également des décalages géométriques (dus à la calibration du capteur).

De façon générale, un certain nombre de facteurs physiques peuvent être à l'origine de l'apparition de défauts dès la phase d'acquisition des données de télédétection. Une série de prétraitements est alors nécessaire avant de pouvoir utiliser les données. Citons quelques exemples types de cas de figures :

- pour garantir la qualité de la détection de changements, les images étudiées doivent être superposables (au niveau du pixel). Ce prétraitement, essentiel à la détection de changements, est appelé recalage d'images. Cependant, des erreurs locales résiduelles de recalage demeurent, en général, entre deux images (par exemple en raison de la complexité topologique de la zone), elles nuisent à la robustesse du détecteur de changements. De nombreuses études montrent l'importance de l'influence de la qualité du recalage d'images sur la précision de la détection de changements [Dai et Khorram, 1998]. Les techniques de recalage sont assez complexes, diverses et fortement liées aux types de capteurs utilisés mais nous ne nous attarderons pas sur le sujet.
- Les instants d'observations, souvent espacés de plusieurs mois voire de plusieurs années, ne permettent pas de garantir des conditions d'observations identiques. Des différences

majeures peuvent être observées au niveau de la réflectivité de la scène, par exemple en fonction de la saison (sèche ou humide).

- Les systèmes d'acquisition optiques ou infra-rouges dépendent de l'ensemble des conditions d'acquisitions et, notamment, de l'angle solaire d'éclairement de la scène, de l'angle d'observation de la scène (angle de visée), du type de trajectoire du satellite (trajectoire montante/descendante) lors des différentes prises de vue, de la composition de l'atmosphère lors des observations, *etc.* Les observations issues de tels capteurs nécessitent donc une correction (normalisation) radiométrique (illumination et/ou réflectivité) afin d'étalonner les images entre elles.
- Dans un cadre opérationnel, les acquisitions peuvent provenir de satellites différents et, par conséquent, avoir des caractéristiques différentes. En effet, les informations acquises sont propres au type de capteur utilisé. De plus, l'utilisation de capteurs identiques n'assure pas forcément l'acquisition d'images comparables [Lee, 2003]. En effet, des difficultés peuvent aussi être liées au fait que les capteurs utilisés [Duggin, 1985] ne sont pas de la même génération, problèmes que l'on peut rencontrer lorsque les deux images sont séparées par plusieurs années (vieillesse des capteurs, résolution spatiale, inter-étalonnage, *etc.*).
- Finalement, l'évolution *naturelle* des scènes observées (phénomène lent comme la croissance de la végétation) ne doit pas être confondue avec les changements abrupts (phénomène rapide).

Les prétraitements associés à ces facteurs sont de nature aussi diverse que les facteurs eux-mêmes. Il est donc nécessaire de mettre en place des techniques spécifiques afin de corriger les défauts inhérents à chaque facteur. En général, ces techniques permettent de corriger ces erreurs en moyenne sur l'image mais pas parfaitement en chaque pixel. Notons qu'une technique de détection de changements est d'autant plus fiable qu'elle est robuste à ces facteurs d'erreur. Le prétraitement des images est, à lui seul, un sujet très vaste sur lequel nous ne nous étendrons pas davantage.

Par la suite, nous supposerons que des images géographiquement recalées, radiométriquement corrigées et normalisées sont disponibles et permettent ainsi une comparaison directe. La section 2.3 présente une revue des principales approches ponctuelles proposées dans la littérature pour la détection de changements.

## 2.3. Les approches ponctuelles

Les méthodes pour lesquelles la décision de détection porte, spatialement, sur un pixel constituent la majeure partie des méthodes de détection. De nombreux détecteurs et techniques dérivées ont été mis au point, notamment selon le type de changement à identifier. Dans un premier temps, nous présentons les approches qui reposent sur la transformation des images considérées dans le but de favoriser la détection des changements éventuels. La section 2.3.1 décrit les dif-

férents types de transformations proposés dans la littérature, tels que la différence d'images, les transformations linéaires ou l'analyse du vecteur des changements. La détection des changements nécessite alors la prise d'une décision pour chaque pixel de l'image. La section 2.3.2 présente différentes méthodes utilisées à cet effet, empiriques ou statistiques. Dans un second temps, nous présentons différents types d'approches pour lesquelles la détection de changements est envisagée sous la forme d'un problème de classification (*cf.* Section 2.3.3) avec une classe pour les pixels de changements et une pour ceux sans changements.

### 2.3.1. Les transformations

Les méthodes les plus couramment utilisées reposent sur le calcul d'opérateurs simples de changements tels que la différence normalisée ou le ratio [Rignot et van Zyl, 1993] entre une image d'archive et une image acquise au cours de l'événement d'intérêt, ou après. Cependant, le ratio d'images est plutôt utilisé dans le cas d'images radar pour lesquelles le bruit est multiplicatif (speckle). Quelques uns des indicateurs les plus utilisés dans la littérature sont présentés ici. Un simple seuillage sur l'image de l'indicateur choisi permet ensuite de partitionner l'image en zones de changements *versus* non-changement.

#### Les opérateurs simples

Les méthodes de détection de changements les plus utilisées, et historiquement les plus anciennes, sont basées sur la **différence d'images** du fait de leur simplicité d'interprétation et de mise en oeuvre. Elles consistent à opérer une différence pixel à pixel entre les images  $I_1$  et  $I_2$  acquises aux dates 1 et 2 pour produire une image dite image de différence  $I_D$ , définie par

$$I_D(x) = I_2(x) - I_1(x),$$

et révélatrice des changements survenus entre deux observations successives en un même pixel  $x$ . La distinction entre les changements et les effets dus aux variations des conditions d'acquisition entre les deux dates considérées est souvent difficile dans les images originales. En général, une transformation est alors appliquée aux images originales avant d'opérer la différence. La littérature présente un certain nombre de travaux discutant de la meilleure combinaison de données à utiliser pour discriminer les changements de types de végétation. Par exemple, [Lyon et al., 1998] présentent la différence d'un indice de végétation, le NDVI (Normalized Difference Vegetation Index), comme la meilleure méthode pour contrôler la déforestation. [Yasuoka, 1988] suggère de normaliser les différentes bandes spectrales de manière à les rendre comparables avant d'en faire la différence.

Les pixels pour lesquels la différence  $I_D$  obtenue est supérieure (en valeur absolue) à un certain seuil sont alors susceptibles d'avoir subi des changements, alors que ceux ayant des valeurs proches de zéro sont considérés comme des pixels inchangés. Ces méthodes sont assez sensibles au bruit et, dans le cas de séquences vidéo, aux mouvements importants dans la scène.

[Coppin et Bauer, 1994] proposent une variante de la différence d'images : la *différence d'image normalisée*. Il s'agit de normaliser la différence d'images par la somme de ces mêmes images. Cet opérateur est défini, pour tout pixel  $x$ , par

$$I_{\mathcal{DN}}(x) = \frac{I_{\mathcal{D}}(x)}{I_1(x) + I_2(x)}.$$

Ce détecteur a été développé pour favoriser la distinction de différents types de changements.

Tout aussi simple que la différence d'images, le **ratio d'images** considère le rapport, pixel par pixel, des intensités mesurées sur les deux images  $I_1$  et  $I_2$  d'une même scène afin de faire apparaître les changements (cf. [Rignot et van Zyl, 1993]). Pour tout pixel  $x$ , l'image de ratio est définie par

$$I_{\mathcal{R}}(x) = \frac{I_2(x)}{I_1(x)}.$$

Cet opérateur est utilisé, en particulier, pour analyser les changements à partir d'images radar (SAR) parce qu'il permet de réduire les distorsions dues au speckle (bruit multiplicatif) des deux images considérées [Singh, 1989]. Pour un pixel donné, l'absence de changements est caractérisée par un ratio proche de 1. À l'inverse, plus le ratio est inférieur ou supérieur à 1, plus le pixel est susceptible d'appartenir à une zone de changements.

Afin d'augmenter la dynamique de l'image de ratio, il a récemment été proposé d'utiliser le logarithme du ratio [Bazi et Melgani, 2005]. En particulier, l'histogramme de l'image de log-ratio décrit une plage de valeurs de pixels allant des valeurs négatives aux valeurs positives, où celles proches de zéro représentent les pixels de non-changement et les queues de l'histogramme représentent des changements de réflectivité entre les deux images [Jensen et Cowen, 1997].

D'autres détecteurs plus élaborés, tels que ceux basés sur des mesures de ressemblance, peuvent être utilisés. À titre d'exemple, citons la mesure de corrélation de Kolmogorov-Smirnov ou l'information mutuelle [Thévenaz et Unser, 2000].

## Les transformations linéaires

Les transformations linéaires les plus utilisées pour la détection de changements reposent sur l'analyse en composantes principales (ACP) ou sur l'analyse canonique de corrélations. Par exemple, la transformation *Détection d'Altérations Multivariée* (MAD) (cf. [Hanaizumi et Fujimura, 1992, Nielsen et al., 1998, Nielsen, 2005]) est basée sur l'analyse canonique de corrélations. Elle permet d'organiser les variables originales en deux groupes de combinaisons linéaires de ces variables : les deux premières combinaisons linéaires correspondent aux variables de plus forte corrélation (la corrélation canonique) et les deux secondes sont celles de plus forte corrélation orthogonales aux premières. La détection de changements est alors fondée sur l'analyse des différences entre les variables de ces deux composantes. Ce type d'approches permet d'ordonner les différences au sein de données multivariées tout en préservant leur structure spatiale.

Dans le contexte temporel de l'analyse des changements dans une série d'images multitemporelle, l'*Analyse en Composantes Principales* (ACP) est utilisée comme un outil qui permet d'obtenir une combinaison linéaire de canaux (des dates, des bandes ou des détecteurs) qui concentre l'information de changements [Fung et Le Drew, 1987, Gong, 1993]. Dans un but de visualisation, la représentation colorée de la seconde composante avec les images de référence permet par ailleurs une cartographie efficace des changements au moment de l'acquisition. Divers cas d'application de l'ACP sont présentés par [Inglada, 2001] en fonction des données d'entrée. En particulier, pour la détection entre deux dates au moyen d'images monobandes, on pourrait supposer que les changements apparaissent dans la deuxième composante. En considérant que les pixels sans changement sont liés de manière linéaire, le poids des pixels de changements présents dans l'échantillon introduit une déviation lors de l'estimation de la composante principale. Une amélioration de la qualité de la détection par ACP ensuite proposée par [Wiemker et al., 1997] en introduisant une certaine régularité spatiale sur l'ensemble des pixels de la deuxième composante. Plus précisément, ils suggèrent de diminuer la pondération des pixels appartenant à une zone à forte variance par rapport aux zones homogènes, de normaliser l'ensemble des points situés sur la deuxième composante par la variance locale avant de régulariser cette deuxième composante de manière à éliminer les pixels isolés, dus probablement au bruit ou à des erreurs de recalage.

En ce qui concerne l'utilisation de plusieurs dates, certains travaux considèrent les pixels de changement d'une séquence d'images comme des pixels de bruit par rapport aux pixels de non-changement. Ces derniers doivent apparaître dans les dernières composantes principales. Cependant, la notion de bruit est très relative à la quantité de pixels de changement dans l'image. Face au manque de robustesse de cette approche, d'autres travaux proposent d'appliquer l'ACP à la différence d'images (cf. [Fung, 1990]), afin d'obtenir les pixels de changements sur la première composante. La différence d'images peut être considérée terme à terme entre dates consécutives ou relativement à la première date.

D'autres méthodes de transformations linéaires ont été proposées dans la littérature, notamment la transformation MAF (Maximum Autocorrelation Factor) capable d'utiliser l'information spatiale pour favoriser la détection de zones étendues (cf. section 2.4.1). En revanche, d'après [Coppin et al., 2004], les transformations non-linéaires n'ont pas encore prouvé leur utilité pour la détection de changements.

### **L'analyse vectorielle des changements**

L'analyse vectorielle des changements est très proche d'une analyse de la différence d'images qui serait réalisée en plusieurs dimensions (spectrale ou temporelle). L'analyse vectorielle est très utilisée pour comparer la différence de trajectoire temporelle d'une intensité (par exemple, une mesure multispectrale, un indice biophysique, une mesure d'hétérogénéité, etc) entre deux périodes successives. Cette trajectoire temporelle est représentée sous la forme d'un vecteur multitemporel  $(V(x, t))_{x, t}$  déterminé en calculant l'indicateur choisi en chaque pixel  $x$ , pour toutes les dates  $t$  de la séquence. L'indicateur considéré est multidimensionnel dans le cas où

plusieurs bandes spectrales ou indices biophysiques sont utilisés, ou monodimensionnel lorsqu'un seul indice ou canal est utilisé pour la détection (cf. [Malila, 1980, Lambin et Stralher, 1994]). Le vecteur de changements correspond alors simplement au vecteur de la différence entre la trajectoire temporelle correspondant à la période étudiée et la trajectoire de référence,  $c(x, t) = V(x, t_{\text{ref}}) - V(x, t)$ , en chaque pixel  $x$ . Il est généralement utilisé pour son intensité qui traduit l'amplitude des changements mais sa direction peut aussi être exploitée pour discriminer les différents types de changements. L'approche couramment adoptée consiste à considérer que le module du vecteur de changements  $c$  est d'autant plus élevé que les changements sont conséquents (cf. [Borak et al., 2000, Bruzzone et Prieto, 2002]). Certains travaux proposent d'appliquer des normalisations ou pondérations au vecteur de changements de façon à rendre les indicateurs comparables ou à prendre en compte un facteur de confiance. Dans le cas le plus général, une norme quelconque peut être utilisée. D'autres approches proposent d'associer la valeur des changements dans chaque composante à une valeur d'appartenance à l'ensemble flou des changements et d'appliquer ensuite des méthodes de fusion classiques en logique floue ou en théorie des croyances de Dempster-Shafer ( cf. [Le Hégarat-Masclé et Seltz, 2004]).

L'*analyse vectorielle des changements* permet donc de traiter des informations de différentes sources indifféremment dans un cadre multitemporel. Cette approche multitemporelle est cependant très sensible aux variations saisonnières, ainsi qu'aux bruits atmosphériques ou bruits de capteurs qui affectent des périodes isolées.

Quel que soit le détecteur considéré, la détection des changements nécessite la prise d'une décision en chaque pixel. Pour ce faire, la section 2.3.2 décrit différentes approches.

### 2.3.2. La décision

Les transformations présentées précédemment permettent la création d'une image dans laquelle les changements apparaissent de manière d'autant plus marquée qu'ils sont importants. Cependant, il est alors nécessaire de décider d'un seuil à partir duquel considérer qu'un changement a eu lieu. Dans cet objectif, certains travaux proposent de fixer un seuil de manière empirique alors que d'autres se placent dans un contexte statistique. Le choix du seuil de décision est déterminant pour la qualité de la détection.

#### Seuillage empirique

A partir d'une image de différence ou de ratio, le seuillage de l'image d'indicateur est, dans la plupart des cas, la technique retenue pour l'obtention d'un masque de changements. Un simple seuillage est effectivement approprié si la zone de recouvrement des histogrammes des classes de changements et de non-changement n'est pas trop grande et lorsque le seuil (valeur limite à partir de laquelle considérer qu'un changement a effectivement eu lieu) est bien choisi. Cependant, ce choix est délicat, il dépend nettement de l'application. Le seuil est souvent choisi empiriquement, même si de nombreux travaux [Smits et Annoni, 2000] proposent une manière de



le déterminer dans un cadre général. Des approches ont été proposées pour le choix automatique d'un seuil global à partir de différents modèles statistiques de l'image [Rosin, 2002]. Ces travaux sont motivés par l'idée qu'un système de vision robuste et autonome ne peut reposer sur le choix expérimental d'un tel seuil. Les méthodes bayésiennes non-supervisées dédiées à l'estimation de paramètres permettent également de discriminer les classes de changements/non-changement à partir de l'histogramme des images de ratio ou de différence.

De plus, indépendamment du problème du choix du seuil, les opérateurs de différence et de ratio sont généralement très sensibles au bruit et aux variations d'illumination. Ils nécessitent, par conséquent, un traitement préalable pour diminuer le bruit dans les images, *i.e.* pour diminuer le recouvrement des histogrammes des classes de changements et de non-changement.

## Tests d'hypothèses

La règle de décision de nombreuses méthodes de détection de changements est basée sur un *test d'hypothèse statistique* classique [Kay, 1993] où l'hypothèse nulle  $\mathcal{H}_0$  : "le pixel  $x$  ne correspond pas à un changement", et l'hypothèse alternative  $\mathcal{H}_1$  serait du type : "le pixel  $x$  correspond à un changement". Une paire d'images est alors modélisée comme un champ de vecteurs aléatoires  $(u_1(x), u_2(x))$ , en chaque pixel  $x$ , pour lesquelles la connaissance de la loi conditionnelle  $p(u_1(x), u_2(x)|H_0)$  permet de choisir l'hypothèse qui décrit le mieux le changement d'intensité en  $x$  en utilisant le cadre classique des tests d'hypothèses. Les méthodes classiques de prise de décision par tests d'hypothèses pour la détection sont décrites dans [Konrad, 2000].

En général, l'hypothèse nulle  $\mathcal{H}_0$  peut être décrite simplement puisque, en l'absence de changements, la différence d'intensité entre les images peut être supposée résultant uniquement du bruit. Typiquement, sous hypothèse  $\mathcal{H}_0$ , la différence peut être considérée sous la forme d'un bruit gaussien de moyenne nulle et de variance  $\sigma^2$ , dépendant du système imageur. Un test de significativité sur l'image des différences permet alors d'évaluer à quel point cette hypothèse décrit l'observation en comparant la probabilité d'observer cette différence sous l'hypothèse  $\mathcal{H}_0$  au seuil de significativité. Ce seuil peut être calculé, *a posteriori*, de manière à garantir, par exemple, un taux minimal de fausses alarmes [Hommel, 1988] ou de fausses découvertes [Benjamini et Hochberg, 1995]. Si la définition de l'hypothèse alternative  $\mathcal{H}_1$  est envisageable, les tests sur le rapport de vraisemblance permettent de détecter les changements en choisissant l'hypothèse qui maximise la probabilité *a posteriori* d'obtenir ces observations. Si l'on suppose par exemple que la différence d'intensité suit une loi uniforme  $[-L, L]$  sous l'hypothèse de changement  $\mathcal{H}_1$ , on peut en déduire le seuil,  $2\sigma^2 \ln \frac{2L}{\sqrt{2\pi\sigma^2}}$ , sur le carré de la différence d'intensité [Konrad, 2000]. Cependant, la caractérisation de  $\mathcal{H}_1$  est généralement difficile en pratique, les changements n'étant pas connus *a priori* ou, s'ils le sont (applications spécifiques), ils sont difficiles à modéliser. Or la forme et les paramètres des lois de probabilités jouent un rôle crucial pour les décisions.

## Règle de décision bayésienne

La règle de Bayes permet de choisir le seuil optimal, au sens de la minimisation du coût moyen, qui sépare deux classes en connaissant leurs distributions de probabilités et en fonction du coût associé aux erreurs. La difficulté réside alors dans l'estimation de ces lois de probabilités.

[Bruzzone et Prieto, 2000] proposent une méthode de sélection automatique d'un seuil de décision de manière à minimiser la probabilité d'erreurs de détection à l'échelle globale sous l'hypothèse selon laquelle les pixels de l'image de différence sont indépendants les uns des autres. L'image de différence  $I_{\mathcal{D}}$  considérée peut correspondre à la différence entre deux dates ou deux vecteurs de changements (pour les images multi-bandes ou les séquences temporelles). Pour chaque pixel, l'objectif est de déterminer s'il appartient à l'ensemble des pixels représentant des changements significatifs,  $\omega_c$ , ou de ceux qui ne représentent pas de changements significatifs,  $\omega_n$ . La probabilité d'observer une différence d'intensité  $I_{\mathcal{D}}$ , de densité  $p(I_{\mathcal{D}})$ , peut alors être décomposée de la manière suivante.

$$p(I_{\mathcal{D}}) = p(I_{\mathcal{D}}|\omega_c)P(\omega_c) + p(I_{\mathcal{D}}|\omega_n)P(\omega_n). \quad (2.3.1)$$

L'algorithme EM est alors utilisé pour estimer les densités  $p(I_{\mathcal{D}}|\omega_c)$ ,  $p(I_{\mathcal{D}}|\omega_n)$  et les probabilités  $P(\omega_c)$  et  $P(\omega_n)$ . Il est initialisé à partir des pixels de valeurs extrêmes dans l'image de différence  $I_{\mathcal{D}}$ . Sous hypothèse d'indépendance spatiale, le seuil optimal de décision qui sépare les deux classes  $\omega_c$  et  $\omega_n$  correspond à la valeur pour laquelle les probabilités de changement/non-changement sont égales :

$$\frac{P(c)}{P(nc)} = \frac{p(I_{\mathcal{D}}|nc)}{p(I_{\mathcal{D}}|c)}.$$

Dans le cas envisagé où les probabilités conditionnelles sont supposées suivre des lois gaussiennes, le seuil de décision peut être obtenu par résolution d'une équation quadratique simple.

Dans ce même article, les auteurs proposent une autre méthode d'estimation du seuil de décision en se plaçant dans un cadre markovien, de manière à considérer la dépendance spatiale entre les pixels (*cf.* section 2.4.2), énergie que l'on peut souvent réinterpréter dans un cadre bayésien. Dans le même esprit, [Moser et Serpico, 2006] proposent une méthode de détection automatique d'un seuil dédiée au cas d'images SAR, *i.e.* en considérant l'image de ratio plutôt que l'image de différence.

### 2.3.3. Les méthodes à base de classifications

Le problème de la détection de changements peut être envisagé comme un problème de classification (non supervisée), où il s'agit de décider si chaque pixel considéré représente un changement ou non.

## La comparaison post-classification

La comparaison post-classification fait partie des techniques les plus utilisées pour détecter les changements entre deux dates ou deux périodes données. La façon la plus simple de procéder consiste à réaliser la classification de la scène étudiée aux deux dates considérées de manière indépendante. Les changements apparus entre deux instants peuvent être identifiés en recherchant les pixels qui ont changé de classe d'une date à l'autre. La génération de la carte de changements s'effectue donc en comparant simplement les deux classifications thématiques obtenues.

Les images de classes peuvent être obtenues en utilisant diverses techniques de classification (déterministe, probabiliste, floue, etc), monodimensionnelles ou multidimensionnelles. Les méthodes de classification multidimensionnelles peuvent être employées pour tenter d'améliorer la qualité de la classification en utilisant plusieurs bandes spectrales à une date donnée, plusieurs indices complémentaires (cf. 2.3.1), ou encore dans un contexte multitemporel pour la comparaison de séquences d'images. La comparaison des classifications obtenues peut ensuite être réalisée de différentes manières, en utilisant des algorithmes de décision ou de comparaison de l'information.

Le traitement indépendant des phases de classification et de détection présente à la fois des avantages et des inconvénients. Parmi les avantages, le fait que les deux acquisitions soient classifiées indépendamment peut laisser penser que les problèmes de calibration et de sensibilité aux conditions atmosphériques entre les dates seront nettement limités. De plus, la détection peut alors être optimisée par le choix d'une méthode de classification appropriée en fonction de l'application. Parmi les inconvénients, la qualité de la détection dépend alors largement de la qualité des classifications à comparer. En effet, la précision de la détection est fortement dépendante de la précision de chaque classification obtenue : un pixel peut être attribué à la classe  $A$  sur l'image  $I_1$ , puis à la classe  $B$  sur l'image  $I_2$ , et ceci pour une faible variation de sa probabilité d'appartenance à la classe  $A$  ou  $B$  ; un tel changement de classe ne permet pas d'affirmer que le pixel observé a effectivement changé. De plus, avec cette méthode, les classifications sont réalisées indépendamment ce qui implique que le nombre de classes n'est pas nécessairement identique aux deux instants ou périodes considérés.

## La classification conjointe

La détection de changements peut aussi être envisagée sous la forme d'un réel problème de classification, pour lequel seules les classes changements et non-changements sont considérées. Dans cet esprit, la classification conjointe consiste à générer la carte des changements à partir d'une méthode de classification conjointe (bidimensionnelle) des deux images originales, ou d'une méthode multitemporelle pour comparer deux périodes de temps. Ce type de méthode présente l'avantage principal de ne pas considérer les changements en termes de pixels, mais plutôt en terme de classes thématiques, c'est-à-dire en tant que modifications spatiales et temporelles des distributions des classes de l'image. En effet, cette stratégie consiste à estimer, au travers d'un algorithme de classification vectoriel, les densités de probabilités bidimensionnelles associées à

chacune des classes dans les deux images. La détection et la localisation des changements peut alors se faire en étudiant et en comparant les distributions marginales bidimensionnelles. Ainsi, une classe ayant des lois marginales différentes en termes de forme est susceptible de montrer une évolution entre les deux acquisitions, et donc de caractériser un changement. Un algorithme de classification pseudo-conjointe basé sur la théorie des ensembles flous a par ailleurs récemment été proposé dans [Agouris et al., 2000]. Tout comme la comparaison post-classification, une difficulté majeure associée à cette stratégie est le choix du nombre de classes pour la classification vectorielle.

### La fusion de classifications

Quelques techniques récentes [Eklund et al., 2000, Deer et Eklund, 2002] tentent d'outrepasser les limites obtenues lors de la comparaison de classifications en proposant d'effectuer la classification initiale des images par des méthodes floues (avec un algorithme de type  $k$ -moyennes floues). La prise en compte du flou dans la classification des images  $I_1$  et  $I_2$  permet alors d'améliorer la qualité de la détection de changements. Une difficulté majeure associée à ce type de méthodes demeure le choix du nombre de classes utilisé pour les classifications initiales. Ce choix requiert des connaissances *a priori* sur les images, connaissances rarement disponibles dans un contexte opérationnel.

Avec une démarche un peu différente, [Le Hégarat-Masclé et Seltz, 2004] proposent d'améliorer la détection de changements par comparaison de classifications mono-indices par la fusion des résultats de détection de changements réalisés à partir de différents indices complémentaires. Cette fusion est réalisée dans le cadre de la théorie de l'évidence de manière à favoriser la prise en compte de l'imprécision de l'information ou de l'ignorance pour la décision entre classes de changements ou de non-changements.

Les masques de détection obtenus en sortie des méthodes de détection ponctuelles (analyse pixel à pixel) ne sont pas toujours satisfaisants. Ces méthodes souffrent généralement d'une probabilité d'erreurs assez élevée : de petites zones de changements ont tendance à être détectées essentiellement à cause du bruit. Du point de vue des applications, la détection de grandes régions connexes est souvent souhaitable. Dans cet objectif, certains travaux proposent simplement d'appliquer des filtres morphologiques pour lisser le masque de détection, mais le masque obtenu a alors tendance à manquer de précision. La section 2.4 présente d'autres approches fondées sur la prise en compte de l'information spatiale des images pour la prise de décision.

## 2.4. Les approches contextuelles

Dans cette section, nous présentons les différents types de méthodes proposées dans la littérature pour prendre en compte l'information spatiale des données images en vue de renforcer la cohérence spatiale du masque de détection. Elles reposent sur l'hypothèse *a priori* selon laquelle les changements sont spatialement étendus. Pour la détection des changements, l'information spatiale

peut être utilisée essentiellement de trois manières : lors de la transformation des données, lors de la prise de décision en se plaçant dans un cadre markovien (*cf.* 2.4.2) ou encore en utilisant des méthodes de classification apparentées région.

### 2.4.1. Les transformations

Diverses méthodes ont été proposées, dans la littérature, afin de prendre en compte l'information spatiale dans la définition des *opérateurs simples* décrits dans la section 2.3.1. Certaines consistent à appliquer les opérateurs, non plus pixel à pixel, mais sur les moyennes calculées dans une fenêtre de régularisation centrée sur le pixel d'intérêt. La prise en compte de ce voisinage permet de rendre les résultats de classification plus homogènes. À titre d'exemple, l'utilisation d'une fenêtre de régularisation pour le calcul du détecteur du ratio permet de réduire considérablement l'influence du chatoiement sur la carte de changements de données radar SAR (*cf.* [Weydahl, 1991]). De façon générale, les fenêtres carrées sont couramment utilisées, leur taille est déterminée de manière empirique, habituellement de  $3 \times 3$  à  $15 \times 15$  pixels. Par ailleurs, d'autres transformations ont été proposées dans le but d'intégrer l'information spatiale avant la prise de décision, telles que le calcul de paramètres texturels ou la transformation MAF.

#### Les paramètres de texture

L'information textuelle peut constituer une information discriminante pour la détection de changements. Les méthodes de caractérisation des textures sont variées et utilisent généralement des caractéristiques spatiales. Certaines constituent en fait une modélisation du signal de texture alors que la plupart des méthodes mesurent simplement quelques attributs discriminants vis-à-vis du problème à résoudre. En pratique, les paramètres obtenus à partir des matrices de co-occurrence, caractérisant la distribution des niveaux de gris sur différents  $n$ -uplets de pixels voisins, ainsi que les moments statistiques, caractérisant la forme de l'histogramme des niveaux de gris à l'intérieur d'une fenêtre, sont les plus utilisés. Généralement, les attributs texturels sont exploités conjointement aux attributs spectraux pour la détection de changements, plus rarement seuls [Smits et Annoni, 2000]. En particulier, ils sont très efficaces pour la détection de changements entre les surfaces de végétation et de sol nu.

#### La transformation MAF

La transformation MAF (Facteurs d'Autocorrélation Minimum/Maximum) peut être considérée comme une extension spatiale de l'analyse en composantes principales (ACP) [Nielsen et al., 1998]. Au lieu de maximiser la variance des composantes, le but de la transformation MAF est de déterminer les composantes qui maximisent l'autocorrélation entre pixels voisins. Cette transformation est constituée de deux composantes : MAF1 et MAF2, où MAF1 correspond à la combinaison linéaire des variables originales qui maximisent l'autocorrélation et MAF2 est alors la combinaison linéaire

des variables originales qui maximisent l'autocorrélation tout en étant orthogonales à MAF1. Cette transformation est invariante par transformation affine des variables originales (corrections radiométriques ou atmosphériques linéaires, etc).

### 2.4.2. La décision

Comme nous l'avons dit précédemment, les techniques basées sur le calcul d'indices pour chaque pixel sont très sensibles au bruit. Une réduction des effets du bruit peut être obtenue par lissage des images avant détection ou par lissage des produits de détection. Le problème de ces approches est que l'on est contraint de définir un voisinage ou une réponse impulsionnelle du filtre qui peut engendrer des problèmes aux frontières des zones de changements. Cette difficulté apparaît aussi dans le choix des voisinages pour l'estimation des textures ou des moments statistiques.

L'approche proposée dans [Bruzzone et Prieto, 2000] consiste à segmenter les deux images (avant et après) et fusionner les segmentations dans le but d'obtenir une segmentation commune aux deux images où chaque parcelle est homogène (par rapport aux critères de segmentation). Ensuite, chaque parcelle est caractérisée par un vecteur d'indices de changements qui permet d'évaluer les changements en restant à la résolution des parcelles. Les résultats obtenus montrent des zones de changements compactes et une faible sensibilité au bruit. Évidemment, cette technique ne peut s'appliquer que dans les cas où la résolution de l'image est suffisamment haute relativement à la taille spatiale des changements à détecter.

Les modèles markoviens sont bien adaptés à la prise en compte d'un contexte spatial, voire temporel, pour la détection de changements. En particulier, dans le cas de l'étude des changements à partir de données radar, quelques travaux très récents ont permis de montrer que le problème de la détection de changements pouvait être traité avec succès par des approches markoviennes [Kasetkasem et Varshney, 2002, Carincotte et al., 2004].

Dans le cadre de l'estimation bayésienne d'un seuil optimal de décision (*cf.* (2.3.1)) il est possible d'utiliser une modélisation par champs de Markov afin d'introduire une dépendance spatiale entre les pixels (ce qui est adapté au cas de capteurs à haute résolution par rapport à la taille des zones de changements) [Bruzzone et Prieto, 2000]. L'approche markovienne permet d'introduire la dépendance de la valeur d'un pixel en fonction de celles de ses voisins. On utilise alors, de même qu'en section 2.3.2, les probabilités calculées par l'algorithme EM et on procède par la méthode classique fournie par le cadre markovien : définition de potentiels, puis minimisation de ceux-ci. Si l'on compare cette approche avec la méthode non-markovienne, on retrouve des résultats similaires, mais la première (indépendance spatiale) est plus sensible au bruit et les zones détectées apparaissent fragmentées. La prise en compte d'un contexte spatial permet en effet d'adapter le seuil de détection en fonction du voisinage d'un pixel.

Parmi les techniques de détection portant sur les régions, on recense également quelques méthodes variationnelles ou méthodes de contours actifs. Ces méthodes reposent sur la minimisation d'une énergie composée d'un terme d'attache aux données et d'un terme de régularisation.

### 2.4.3. Les méthodes à base de classifications

L'approche par *classification conjointe* consiste à générer la carte de changements à partir de la classification conjointe des deux images originales, à l'aide d'un algorithme statistique de classification. Elle présente l'avantage de ne pas considérer les changements en terme de pixels, mais plutôt en termes de classes thématiques, c'est-à-dire en tant que modifications spatiales et temporelles des distributions des classes de l'image. En effet, cette stratégie vise à estimer, au travers d'un algorithme de classification vectoriel, les densités de probabilités bidimensionnelles associées à chacune des classes dans les deux images. La détection et la localisation des changements peut alors se faire en étudiant et en comparant les distributions marginales bidimensionnelles. Ainsi, une classe ayant des lois marginales différentes en termes de formes est susceptible de montrer une évolution entre les deux acquisitions et, donc, de caractériser un changement. Un algorithme de classification pseudo-conjointe basé sur la théorie des ensembles flous a par ailleurs récemment été proposé par [Agouris et al., 2000]. Tout comme la comparaison post-classification, une difficulté majeure associée à cette stratégie vient du choix du nombre de classes dans la classification vectorielle.

Comme nous l'avons remarqué dans la section précédente, la méthode de *comparaison post-classification* repose sur la comparaison de deux classifications d'une même scène à deux dates distinctes. Plus précisément, s'il s'agit de comparer une zone géographique à deux intervalles de temps, une classification est réalisée à l'issue de chaque période de temps. Afin d'introduire de l'information de nature spatiale, la comparaison des classifications peut être réalisée pixel par pixel, telle que décrite dans la section 2.3.3, ou région par région pour détecter, par exemple, les changements de type de couvert. Cependant, les erreurs de classification sont fréquentes et contribuent à limiter les performances de ce type de méthodes.

## 2.5. Conclusion

Au cours de ce chapitre, nous avons présenté un certain nombre de méthodes (*cf.* [Singh, 1989, Nielsen et al., 1998, Mas, 1999, Smits et Annoni, 2000, Bruzzone et Prieto, 2002, Le Hégarat-Masclé et Seltz, 2004]) développées pour détecter des changements sur la surface terrestre à partir d'images haute résolution, c'est-à-dire avec des pixels couvrant une surface inférieure à  $30 \times 30 \text{ m}^2$  (**Spot/HRV, Landsat/TM**) ou, plus généralement, avec des images dont les pixels sont homogènes (pixels purs, *i.e.* correspondant à une seule classe d'objets). Pour certaines études plus globales, à échelle nationale ou planétaire, les objets d'intérêt peuvent être spatialement *grossiers* (forêt, zones agricoles, océans) et les pixels des images grand champ (basse résolution spatiale) sont alors homogènes à l'échelle de ces objets. Dans ce cas, les méthodes présentées dans ce chapitre peuvent être appliquées à des images basse résolution. Par exemple, les méthodes fondées sur l'analyse du vecteur de changement sont communément utilisées (*cf.* [Lambin et Stralher, 1994, Borak et al., 2000]) pour des applications de suivi de la végétation à l'échelle globale (à partir d'images basse résolution).

Cependant, si les études centrées sur l'analyse et le suivi de la végétation nécessitent l'utilisation de données grand champ pour leur haute répétitivité (permettant l'acquisition des profils temporels des différentes classes), elles nécessitent aussi une bonne localisation des objets d'intérêt. Pour des applications liées à la cartographie ou à l'inventaire des cultures, les objets géographiques tels que les parcelles agricoles ou les fleuves sont souvent d'une taille inférieure à celle de la surface observée à l'échelle d'un pixel d'un capteur grand champ. Le développement de méthodes capables de détecter des changements de taille inférieure à la taille du pixel considéré est alors nécessaire.

La détection des changements sous-pixelliques peut être vue comme un problème de détection couplé à un problème de désagrégation des pixels mixtes ou de super-résolution. Jusqu'à présent, cette problématique n'a pas été envisagée directement. En revanche, de nombreuses approches ont été proposées pour la désagrégation des pixels mixtes ou pour l'amélioration de la résolution (super-résolution). De plus, quelques travaux récents ont proposé des méthodes de classification sous-pixelliques. Dans le chapitre 3, nous présentons une revue de l'ensemble des méthodes d'analyse sous-pixellique.





## 3. Méthodes d'analyse sous-pixellique

Les intensités mesurées en chaque pixel par les capteurs satellites résultent souvent d'un mélange spectral dû aux réponses des différentes classes présentes sur la région observée. Dans ce chapitre, nous présentons une revue des différents types de méthodes reportées dans la littérature et permettant de traiter les images à une échelle sous-pixellique.

### 3.1. Motivation

L'analyse de la surface terrestre et, en particulier, du couvert végétal dépend de l'échelle d'observation (*i.e.* de la résolution spatiale du capteur utilisé) relativement à la taille des éléments d'intérêt. Dans le chapitre 1, nous avons souligné l'intérêt des images satellites à forte répétitivité temporelle malgré leur résolution spatiale souvent plus grossière que l'étendue des parcelles ou motifs géographiques présents au sol. Le facteur d'échelle entre les données et les objets étudiés conduit alors à la présence de pixels **mixtes** dans les images, dans le sens où ils correspondent à une réponse spectrale composite résultant du mélange des réponses spectrales des différents types de couvert présents au sein du pixel. C'est le cas notamment des séquences d'images acquises avec une forte répétitivité et une résolution spatiale grossière. Le mélange des réponses spectrales de pixels frontaliers peut aussi conduire à la présence de pixels mixtes.

L'un des paramètres fondamentaux pour le contrôle de la végétation et l'évaluation des systèmes agricoles est l'identification des différents types de culture. Par exemple, les motifs spatiaux d'un paysage représentent une variable clef pour la gestion et la compréhension de l'environnement. L'accès à des informations sous-pixelliques est donc essentiel pour ce type d'applications (cartographie, suivi des zones urbaines ou des inondations, etc).

Décrites dans le chapitre 2, les méthodes proposées pour la détection de changements à partir de données de télédétection permettent d'extraire l'information au niveau du pixel. Dès que le champ de vision d'un pixel est plus grand que les éléments d'intérêt au sol, chaque pixel couvre plus d'un élément géographique (parcelle, plan d'eau, route, etc). Les méthodes d'analyse à l'échelle du pixel sont alors incapables d'accéder aux informations liées à ces éléments. De manière générale, nous qualifions d'analyse sous-pixellique une analyse permettant d'accéder à des informations concernant des éléments de taille inférieure à celle du champ de vision du pixel du capteur considéré.

Dans ce chapitre, nous présentons les différents types de méthodes proposées pour l'analyse sous-pixellique, appelée aussi analyse super-résolution dans le sens où il s'agit d'accéder à une information de résolution plus fine que la résolution initiale des données exploitées. De nombreuses méthodes de fusion de données multirésolutions ont été proposées dans la littérature, dédiées en général à de faibles rapports de résolution (2 ou 4). Elles permettent de reconstruire des images à résolution plus fine et de se ramener à un problème d'analyse classique à la nouvelle résolution du pixel reconstruit. La section 3.2 présente les principaux types d'approches rencontrés (pyramides morphologiques, décomposition en ondelettes, *etc*). Par ailleurs, face au mélange de l'information, les méthodes les plus répandues pour l'analyse sous-pixellique d'images de télédétection reposent sur la désagrégation de l'information. Il s'agit de décomposer la mesure spectrale acquise en un pixel en fonction des différentes classes d'objets représentées au sein du pixel. La section 3.3 présente les principaux travaux réalisés dans cet objectif. Plus récemment, d'autres travaux se sont intéressés à la classification sous-pixellique dans le but d'identifier les différents types d'occupation du sol avec une résolution plus fine que celle des images exploitées. La section 3.4 en présente une vue d'ensemble.

### 3.2. Les méthodes de fusion de données multi-résolutions

L'objectif des méthodes de fusion multirésolution est de reconstruire des images à haute résolution en préservant l'essentiel de l'information spatiale provenant de l'image haute résolution et de l'information spectrale de l'image basse résolution. Ce type de méthodes permet, notamment, d'améliorer les résultats de classification ou d'autres traitements. L'intérêt de la combinaison d'informations de résolutions différentes vient généralement du fait que les images basse résolution fournissent des informations sur les régions homogènes alors que les images haute résolution permettent de distinguer les contours principaux et les bords des structures importantes. La combinaison des deux sources d'information est réalisée dans l'objectif de réduire la complexité des données tout en conservant les informations complémentaires qu'elles fournissent.

De nombreuses approches ont été proposées dans la littérature, en particulier fusionner des données de type panchromatique avec des données multispectrales. Dans cette section, nous présentons succinctement les méthodes les plus utilisées en télédétection. Un état de l'art sur les nombreux problèmes posés par la fusion de données en général, ainsi que les diverses méthodes proposées pour les résoudre, est détaillé dans [Pohl et Van Genderen, 1998]. Les techniques développées varient d'une application à une autre et s'étendent du simple schéma de fusion par moyennage des pixels à des schémas plus complexes tels que ceux issus de transformées en ondelettes [Ranchin et Wald, 2000].

Certains travaux proposent simplement de combiner les intensités mesurées par les différentes données. Par exemple, la combinaison arithmétique de bandes spectrales ou panchromatiques haute et basse résolution est souvent utilisée dans un but d'amélioration visuelle des données. Il s'agit simplement, après ré-échantillonnage des données à la même taille, de mélanger les deux types

d'images par addition, soustraction ou produit (*cf.* [Carper et al., 1990, Chavez et Anderson, 1991, Filiberti et al., 1994]). Cependant, l'objectif de ce type d'approche est souvent davantage de créer des images satisfaisantes visuellement que de reconstruire l'information spectrale avec fiabilité. Si la combinaison opérée conduit à des résultats satisfaisants visuellement, elle risque d'engendrer de fortes distorsions spectrales.

Dans l'objectif d'amoindrir ces distorsions, de nombreux travaux ont donné lieu à des méthodes de changements d'échelle permettant de traiter séparément les informations spectrales et spatiales. Généralement fondées sur la décomposition en ondelettes [Sveinsson et Benediktsson, 2000, Tseng et al., 2001, Ranchin et Wald, 2000] ou sur la décomposition pyramidale [Aiazzi et al., 1999, Laporterie-Déjean et al., 2003], elles offrent la possibilité d'extraire les éléments à la résolution spatiale adéquate. Par exemple, [Ranchin et Wald, 2000] utilisent la transformée en ondelettes pour injecter l'information structurelle extraite de l'image haute résolution dans le canal basse résolution (méthode ARSIS : Augmentation de la Résolution Spatiale par Injection des Structures). La transformation dans une base d'ondelettes permet de décomposer l'image, à un niveau de résolution donné, en une image de détails et une approximation. Le principe de la méthode ARSIS est de remplacer les approximations de l'image haute résolution par l'image basse résolution, puis de reconstruire l'image à l'aide des détails de l'image haute résolution.

Dans le même esprit, [Laporterie-Déjean et al., 2003] utilisent les pyramides morphologiques pour extraire les éléments les plus fins à chaque niveau de résolution de la pyramide. Les distorsions spectrales sont alors amoindries mais ce type de méthode présente de fortes limites lorsque le facteur de résolution augmente. A l'exception de l'approche de la méthode de [Laporterie-Déjean et al., 2003] qui permet de fusionner des données multispectrales et grand champ dans un rapport de résolution supérieur à 10, les méthodes de fusion ne sont performantes que dans un rapport de résolution inférieur ou égal à 4.

Indépendamment de la capacité de ces méthodes à traiter de rapports de résolution élevés et à préserver la qualité spectrale des informations, les méthodes de fusion peuvent être considérées en tant que prétraitement pour créer un jeu de données haute résolution à partir des images basse résolution. Une méthode de détection de changements telle que celles présentées dans le chapitre 2 est alors directement utilisable. Cependant, ces méthodes s'appliquent généralement à des images haute résolution dans un contexte bi-date ou à des images basse résolution dans un contexte multitemporel (analyse du vecteur de changements). L'utilisation de ces méthodes dans un contexte multitemporel haute résolution risque d'être coûteux en temps de calcul, étant donné la taille des images haute résolution. De plus, il nous semble plus intéressant de réaliser la détection de changements directement à partir d'images basse résolution et d'éviter ainsi d'introduire des distorsions radiométriques sur les données.

### 3.3. La désagrégation de l'information

La désagrégation des pixels mixtes est une méthode très utilisée pour améliorer l'interprétation des données à résolution spatiale grossière, relativement aux éléments d'intérêt. En particulier, cette approche est classique pour la caractérisation des cultures à partir de données grand champ. Les méthodes de désagrégation peuvent être utilisées aussi bien pour des mélanges spectraux que pour des mélanges spatiaux d'information.

Les principales techniques de désagrégation utilisées pour l'analyse sous-pixellique sont les méthodes floues du type fuzzy c-means ou liées à la théorie des possibilités ([Foody, 1996, Foody, 2000]), les méthodes fondées sur des réseaux de neurones ([Binaghi et al., 1999, Foody, 2000]) ou sur les séparateurs à vastes marges (*cf.* [Brown et al., 2000]) et celles qui exploitent le modèle linéaire de mélange ([Settle et Drake, 1993, Uenishi et al., 2005]). Chacune de ces méthodes permet d'estimer les proportions des différents types de couvert présents en chaque pixel.

Les méthodes de désagrégation du type fuzzy c-means ont tendance à faire l'amalgame entre coefficients d'appartenance d'une classe à un pixel et proportion de cette classe dans le pixel. Dans un contexte d'application où la mesure physique est linéaire, l'introduction de l'incertitude semble peu adaptée pour la modélisation du mélange spectral alors que le modèle linéaire de mélange a fait l'objet de validations physiques dans de nombreux cas d'applications.

C'est pourquoi le modèle linéaire de mélange est très utilisé pour l'analyse de la surface (*cf.* section 5.1). Il suppose que l'intensité mesurée en un pixel mixte est égale, en espérance, à la moyenne des intensités caractéristiques de chaque type d'occupation du sol, pondérée par son taux d'occupation au sein du pixel. Ce modèle est utilisé principalement pour deux types d'applications :

- l'estimation des caractéristiques spectrales des classes à partir des mesures de pixels mixtes, connaissant les proportions des classes au sein de chaque pixel (*cf.* [Faivre et Fischer, 1997, Van Der Meer, 1999, Zhukov et al., 1999]) ;
- l'estimation des proportions de chaque classe au sein d'un pixel **BR**, connaissant les caractéristiques spectrales des classes (*cf.* [Cardot et al., 2003]).

Lorsque les caractéristiques des classes sont connues, ce modèle est utilisé pour estimer les proportions de chaque classe au sein d'un pixel basse résolution. [Cardot et al., 2003] réalisent ce type d'estimation en vue de prédire l'état de la surface à partir d'images de télédétection. Cette information décrit la composition de chaque pixel en termes des différents types d'occupation du sol qui y sont présents mais elle n'apporte aucune indication sur la distribution spatiale de ces types au sein du pixel. Pour un ensemble de cartes de proportions données, de multiples classifications haute résolution sont alors plausibles. Ce type de résultat ne permet donc pas directement d'accéder à une classification des types d'occupation du sol. Cependant, certains travaux proposent de générer une classification à partir des images de proportions en déterminant, par exemple, la classe dominante de chaque pixel. Une partie de l'information est alors perdue.

D'autres méthodes (*cf.* [Le Hégarat-Masclé et al., 2005]) sont centrées sur l'estimation jointe des proportions et des moyennes des classes. Lorsque ces deux paramètres sont inconnus, des

méthodes dérivées des algorithmes EM (Expectation-Maximization, cf. [Dempster et al., 1977]) ou de l'algorithme ICE (Iterative Conditional Estimation, cf. [Pieczynski, 1992]) proposent, alternativement, de fixer un des paramètres pour estimer l'autre.

Plus précisément, [Le Hégarat-Masclé et al., 2005] proposent de détecter les changements de types de couvert à partir d'une séquences d'images **BR**, du nombre de classes et de la composition des pixels **BR** à la date de référence (éventuellement déduite d'une classification **HR** antérieure). Dans l'esprit des approches de type EM ou ICE, cette méthode est basée sur l'estimation alternée d'un paramètre caché (la composition des pixels en termes de classes) et de l'inconnue (les caractéristiques des classes) à partir d'un sous-ensemble de pixels de l'image. La méthode fonctionne en deux étapes : la détection de changements et l'estimation des changements. La première étape permet d'estimer les sous-domaines de l'image correspondant aux pixels de changements et de non-changement. A partir de cette première estimation, la distribution des classes au sein d'un pixel **BR** est ré-estimée, puis comparée à la première distribution de manière à fournir une estimation de la détection de changements *a posteriori*. Cette méthode ne peut fonctionner que si les pixels de changements sont minoritaires dans l'image (moins de 30% de l'image). Elle permet, cependant, de détecter des changements dès qu'ils représentent plus de 30% du pixel **BR**. En revanche, elle ne fournit pas d'information quant à la localisation des changements sous-pixelliques dans le pixel impacté.

Mêmes si elles prouvent une certaine efficacité dans la plupart des cas, ces méthodes ne permettent pas de connaître la distribution spatiale des classes au sein du pixel. La désagrégation spectrale est donc très utilisée pour évaluer les taux de couverture du sol par les différents types de végétation. Elle permet de traiter de forts rapports de résolution mais, jusqu'à présent, elle a été utilisée essentiellement pour produire des images de proportions dans lesquelles chaque image représente le pourcentage d'un type d'occupation du sol donné dans chaque pixel sans localisation de ces classes dans le pixel. Ce résultat est parfois qualifié de *classification floue* dans le sens où la localisation spatiale des éléments n'est pas précisée. Cependant, certains travaux ont été proposés récemment pour la classification d'images à une échelle sous-pixellique. Ces méthodes, présentées dans la section 3.4, permettent de considérer la distribution spatiale des classes présentes au sein de chaque pixel de manière à produire des cartes du couvert à une résolution sous-pixellique. Pour cela, elles utilisent un modèle structurel *a priori* où une image supplémentaire acquise avec une résolution plus fine.

### 3.4. La classification sous-pixellique

Les images de télédétection peuvent contenir une proportion significative de pixels représentant un mélange de différents types de couvert, surtout lorsqu'elles sont acquises avec une résolution spatiale grossière. L'analyse des pixels mixtes constitue un problème difficile pour la cartographie du couvert puisque leurs caractéristiques spectrales ne sont pas représentatives d'un seul type de couvert. L'objectif des méthodes de classification sous-pixellique est donc de permettre à

la fois la distinction des différents types d'occupation du sol et leur localisation pour fournir une carte d'occupation des sols à une résolution plus fine que celle des données utilisées. En particulier, la résolution de la classification doit être suffisamment fine pour permettre de distinguer la structure spatiale de la scène observée. Les méthodes de désagrégation présentées dans la section 3.3 peuvent être considérées comme des méthodes de classification floue dans le sens où elles n'apportent pas d'information quant à la localisation des éléments géographiques sous-pixelliques mais simplement les images des proportions en chaque pixel pour tous les types de couvert. Dans cette section, nous qualifions de classification sous-pixellique uniquement les méthodes capables de préciser la localisation des éléments d'intérêt. Dans la littérature, on trouve principalement deux types d'approches : celles qui exploitent l'information de voisinage des images et celles qui utilisent une information complémentaire à une échelle plus fine (images haute résolution ou modèle structurel *a priori*).

Certaines méthodes proposent de subdiviser chaque pixel basse résolution en sous-pixels et de classifier ces derniers en maximisant leur auto-corrélation spatiale (*cf* [Atkinson, 1997, Atkinson, 2001, Verhoeve et De Wulf, 2002, Mertens et al., 2003]). Ces méthodes s'appliquent au résultat préalable d'une classification floue fournissant les images de proportion de chaque type de couvert. Elles reposent sur une hypothèse de dépendance spatiale entre les sous-pixels et pixels, selon laquelle il est plus vraisemblable qu'une même classe soit représentée par des observations voisines plutôt que par des observations séparées.

La première méthode permettant de réaliser une classification sous-pixellique simplement à partir d'une classification floue (sans données supplémentaires) a été proposée par [Atkinson, 1997]. La maximisation de la corrélation spatiale entre pixels voisins se fait alors sous contrainte de conservation des proportions d'origine dans chaque pixel. Si cette méthode a prouvé son efficacité pour des objets de forme simple, elle est limitée dans le cas de données réelles, plus complexes, telles que celles acquises par le capteur **SPOT/HRV**. Un algorithme d'optimisation fondé sur l'échange de pixels a ensuite été proposé par [Atkinson, 2001] pour la mise en œuvre de ce modèle. L'arrangement spatial des valeurs des sous-pixels est modifiée de manière itérative pour maximiser la corrélation spatiale entre sous-pixels voisins. Cet algorithme repose sur la définition d'une fonction de poids exponentielle paramétrée empiriquement. [Makido et Shortridge, 2005] proposent une amélioration de cet algorithme par l'emploi d'un recuit simulé.

[Verhoeve et De Wulf, 2002] envisagent le problème de la classification sous-pixellique sous la forme d'un problème linéaire d'optimisation pour maximiser l'autocorrélation spatiale. Cette méthode permet d'obtenir une classification du couvert sans l'aide de données de résolution plus fine mais elle n'envisage pas l'existence de solutions multiples au problème d'optimisation.

[Mertens et al., 2003] considèrent la même formulation du problème mais ils utilisent un algorithme générique pour la recherche des cartes super-résolution plausibles. [Mertens et al., 2004] utilisent les ondelettes pour relier la différence de résolution entre les classes haute résolution et les proportions basse résolution. À une échelle fine, un réseau de neurones permet d'estimer les coefficients en ondelettes à partir desquels une carte super-résolution est reconstruite.

Ce type d'approches permet d'augmenter artificiellement la résolution spatiale des données bien que cela ne corresponde pas réellement à une augmentation de l'information. Le principe de base de ce type de méthodes est raisonnable tant que les objets d'intérêt sont plus grands que la surface représentée par un pixel ou bien s'ils sont situés simultanément sur deux pixels voisins. En particulier, ce type de méthode ne permet pas de localiser des objets représentés totalement à l'intérieur d'un pixel. Une manière d'accéder réellement à une information sous-pixellique, ciblée en fonction des objets d'intérêt, est d'utiliser des informations haute résolution complémentaires.

Dans le but d'accéder à une information spatiale sous-pixellique, d'autres techniques utilisent une image à résolution spatiale plus fine que celle de l'image à classifier. Dans cet esprit, [Foody, 1998] utilise les contours d'une image haute résolution pour préciser le résultat d'une classification floue dans un contexte de régression. Cependant, les résultats obtenus pour la localisation des objets se sont révélés limités.

La méthode proposée par [Tatem et al., 2002] utilise un modèle *a priori* sur la distribution spatiale des objets pour produire une classification super-résolution à partir d'une classification floue initiale. Elle repose sur l'utilisation d'un réseau de neurones de Hopfield comme outil d'optimisation d'une fonction d'énergie pour prédire le motif spatial des caractéristiques du couvert sous certaines contraintes. La première contrainte est donnée par les images des proportions obtenues par classification floue. La deuxième est incorporée dans la fonction d'énergie sous la forme d'une fonction de semi-variance décrivant l'arrangement spatial *a priori* des différentes classes de couvert. Cette méthode s'inscrit dans la continuité de [Tatem et al., 2001] où l'information *a priori* était réduite à la cohérence spatiale, *i.e.* à la propriété de similitude d'objets voisins.

[Boucher et Kyriakidis, 2006] proposent d'utiliser un modèle structurel *a priori* pour contraindre l'espace des configurations spatiales possibles et présenter l'ensemble des classifications haute résolution plausibles à partir des cartes des proportions obtenues à basse résolution. En effet, les auteurs partent du principe que, même avec un modèle de structure spatiale, plusieurs solutions de classification plausibles subsistent et sont à explorer.

Par ailleurs, [Kasetkasem et al., 2005] proposent d'exploiter le cadre des champs de Markov pour tenir compte de la dépendance spatiale entre les pixels. L'image basse résolution est d'abord utilisée pour générer une classification floue initiale qui est, ensuite, raffinée en tenant compte de la dépendance spatiale des pixels voisins. Si cette méthode permet de réduire significativement le nombre de pixels isolés dans le résultat de classification, les facteurs de résolution traités doivent être suffisamment faibles pour que les objets d'intérêt soient situés, au moins, sur deux pixels voisins.

## 3.5. Conclusion

Dans ce chapitre, nous avons présenté une revue des méthodes d'analyse sous-pixellique rencontrées dans la littérature. Les méthodes de fusion de données multi-résolution sont nombreuses et de plus en plus performantes. Cependant, elles traitent généralement de faibles facteurs de



résolution. De plus, leur utilisation en tant que prétraitement pour la détection de changements ou la classification dans un contexte multitemporel conduirait à de forts volumes de données difficiles à traiter. Sans information complémentaire, les méthodes de désagrégation de l'information ne permettent pas réellement d'obtenir des informations à une échelle sous-pixellique. Cependant, elles permettent de prendre en compte la mixité de la mesure en chaque pixel pour élaborer une classification dite floue, capable de préciser les proportions des différentes classes en chaque pixel sans localisation spatiale. Les méthodes de classification sous-pixellique, développées plus récemment, cherchent à préciser la distribution spatiale des différentes classes sous-pixelliques. Dans cet objectif, certains travaux reposent sur une hypothèse de dépendance spatiale entre pixels et sous-pixels voisins. Dès que l'objet d'intérêt est totalement inclus dans le pixel, la localisation n'est alors plus possible. Par conséquent, d'autres travaux proposent d'exploiter une information complémentaire, de résolution plus fine, pour préciser la localisation des différentes classes. Cette information complémentaire est souvent considérée sous la forme d'une image haute résolution ou un modèle structurel haute résolution.

Parmi ces méthodes d'analyse, très peu de travaux se sont intéressés à la détection de changements. Une comparaison post-classification des classifications sous-pixelliques est parfois réalisée pour détecter les changements, mais cette approche dépend largement des qualités des méthodes de classification utilisées, du nombre de classes et de la bonne correspondance des classes entre les deux classifications. Si les travaux récents de [Le Hégarat-Masclé et al., 2005] ont proposé une méthode de détection des changements sous-pixelliques par désagrégation de l'information et estimation alternée des paramètres de proportions et des caractéristiques des classes, le problème de la détection de changements à une échelle sous-pixellique reste une question ouverte.

## 4. Problématique

Au terme de cette partie, nous avons abordé l'ensemble des éléments qui nous ont conduit à définir la problématique de la détection de changements dans le contexte du suivi de la surface continentale. Si l'information spatiale est importante pour la localisation des changements, l'information temporelle est essentielle non seulement pour la régularité du suivi mais aussi pour permettre la discrimination des différentes espèces de végétation. Actuellement, les capteurs satellites capables de fournir des mesures multispectrales de la surface avec une fine résolution spatiale sont nombreux. Cependant, le compromis entre résolution spatiale et fréquence temporelle est inévitable, intrinsèque aux propriétés physiques des capteurs (*cf.* section 1.2). Cette dualité incite au développement de méthodes capables d'exploiter conjointement des images haute résolution (du type **SPOT/HRVIR**) et des séquences d'images moyenne ou basse résolution (du type **MeRIS** ou **SPOT/VGT**). En effet, les images haute résolution spatiale fourniraient alors l'information spatiale (géométrique) de la surface géographique observée et les séquences d'images basse résolution apporteraient des mesures spectrales ciblées avec une fréquence temporelle adaptée à l'application.

Si des méthodes de détection de changements ont déjà été proposées dans le contexte du suivi de la surface continentale, elles s'inscrivent généralement dans un schéma basse résolution (*cf.* analyse des vecteurs de changements 2.3.1). Dans un schéma multidates, la comparaison de vecteurs de changements peut conduire à des détections simplement parce que le cycle phénologique de certaines espèces a subi un décalage d'une année à l'autre en raison de la variabilité inter-annuelle des saisons. Au-delà des problèmes engendrés par ce type de méthodes pour le choix de la mesure à considérer et du seuil de détection, ces méthodes ne permettent pas de détecter des changements concernant des éléments de taille inférieure à la surface représentée dans un pixel.

La comparaison de classifications sous-pixeliques pourrait être une solution pour la détection des changements lorsqu'ils sont de taille inférieure à celle du pixel basse résolution. Cependant, les méthodes de détection par comparaison de classifications sont généralement très limitées à la fois par la qualité des classifications utilisées et par le fait de comparer des classes qui ne sont pas nécessairement comparables. En effet, lorsque les classifications sont générées indépendamment les unes des autres, les classes obtenues sont différentes à chaque date considérée et les différences entre les classifications peuvent résulter simplement des variations de l'algorithme de classification.

En dehors des approches liées à la classification, les méthodes développées reposent généralement sur la comparaison des valeurs radiométriques (ou dérivées) obtenues en un pixel à deux dates différentes. Elles nécessitent alors un interétalonnage des images avant de pouvoir les comparer et restent, cependant, sensibles aux variations des conditions d'acquisition et, en

particulier, d'illumination.

Face aux problèmes relatifs à la variabilité intrinsèque des images et à ceux posés par les conditions d'acquisitions, nous proposons de détecter les changements par comparaison d'une séquence d'images (ou image) à une classification de référence. La classification joue alors le rôle d'un *masque* permettant d'étudier la cohérence spatiale de la séquence d'images. En effet, cette approche repose alors sur la comparaison entre une valeur obtenue ponctuellement, en un pixel, et une mesure intégrée à l'échelle de l'image qui serait caractéristique de la classe correspondante. Finalement, sous cet angle de vue, l'apparition de changements serait alors détectée si la cohérence interne d'une séquence d'images ne suit pas celle dictée par la classification de référence. Cette démarche permet, en particulier, de nous affranchir des problèmes d'interétalonnage d'une date à l'autre.

D'autre part, compte tenu des propriétés des données et dans l'objectif d'un suivi régulier de la surface continentale, la question de la détection des changements se pose à une échelle sous-pixellique. Pour cela, nous proposons de détecter des changements d'occupation du sol en comparant une séquence d'images **basse résolution** à une classification **haute résolution** de la même scène à la date de référence de manière à être capable de détecter les changements qui apparaissent sur la surface même lorsqu'ils ne touchent qu'une partie du pixel basse résolution.

Dans ce contexte, nous nous intéressons aussi au problème de la classification de la surface continentale, à une échelle sous-pixellique. Par conséquent, dans la partie II, nous abordons le problème de la classification haute résolution à partir d'une séquence d'images basse résolution. En partie III, nous proposons une méthode de détection de changements fondée sur la comparaison d'une séquence d'images basse résolution à une classification haute résolution de référence. Ces deux méthodes pourront alors être utilisées indépendamment ou conjointement, notamment pour la mise à jour de cartographies.

Deuxième partie .

## Classification sous-pixelique de l'occupation du sol



# Introduction

Comme nous l'avons souligné dans la section 1.3, une bonne fréquence d'acquisition de l'information est nécessaire pour caractériser les différents types d'occupation du sol. Compte tenu des caractéristiques des capteurs satellites, un compromis est alors inévitable entre haute résolution spatiale et temporelle. Si la haute résolution spatiale permet une bonne localisation des objets d'intérêt, elle peut ne pas suffire - même couplée à une résolution spectrale adaptée - à la discrimination des différents types de couvert. La haute fréquence temporelle, en revanche, permet de caractériser l'occupation du sol grâce à l'évolution radiométrique temporelle qui est souvent discriminante pour les différents types de végétation. De plus, les images acquises avec une haute fréquence temporelle bénéficient d'un grand champ de vision particulièrement adapté aux études à échelle régionale voire globale, au détriment d'une résolution spatiale moyenne ou grossière (de l'ordre de 300m ou 1km).

L'objectif de cette partie est de développer un outil de classification à partir de séries temporelles d'images grand champ, capable de localiser les objets à une échelle sous-pixellique. La définition suivante précise la notion de classification telle que nous l'entendons.

**Définition 4.0.1 (classification)** *Etant donné un ensemble de classes (représentant une ou plusieurs caractéristiques), une application du domaine de l'image dans l'ensemble des classes qui associe une classe à chaque pixel du domaine est appelée classification. En pratique, l'image des classes qui résulte de cette application est aussi appelée classification.*

Dans ce document, nous distinguons les classes des labels qui leur sont associés : une classe désigne une entité en elle-même, *i.e.* un type d'occupation du sol (ex : forêt, blé, eau, etc.), alors que le label est un numéro attribué pour représenter une classe donnée (ex : 1 pour forêt, 2 pour blé, etc). Avec cette terminologie, une classification est une labellisation couplée à une interprétation thématique des différents labels.

Dans la littérature, les travaux antérieurs proposent des solutions souvent partielles pour la classification sous-pixellique (*cf.* chapitre 3). En effet, les méthodes de classification dites floues ne permettent pas la localisation des différentes classes au sein d'un pixel basse résolution et les méthodes qui reposent sur la prise en compte des interactions spatiales entre pixels voisins (par exemple, avec une hypothèse de champs de Markov) ne sont pas adaptées aux forts rapports de résolution, tels que ceux définis par l'utilisation conjointe des capteurs **SPOT/HRV** et **MeRIS** (rapport de  $15 \times 15$ ) ou **SPOT/HRV** et **SPOT/VGT** (rapport de  $50 \times 50$ ), où la taille des objets d'intérêt est trop petite relativement au champ de vision d'un pixel.

---

L'approche que nous proposons repose sur la théorie bayésienne, avec le critère du *Maximum A Posteriori* et suppose que la structure spatiale de la scène étudiée est invariante au cours de la période considérée. Avec cette hypothèse, l'information spatiale de la scène peut être extraite à un instant donné de la période d'observation, par exemple en segmentant une image **HR**. Une telle hypothèse est réaliste pour une période de temps telle qu'une année agricole, en particulier si l'image **HR** est sur-segmentée de sorte à favoriser la discrimination de toutes les parcelles de la zone d'intérêt.

Néanmoins, cette hypothèse de stationnarité devient contraignante lorsque des changements surviennent au cours de la période d'intérêt. La partie III, centrée sur le problème de la détection de changements, pourra alors être utilisée en collaboration avec la classification, par exemple pour un traitement complet de mise à jour de classification.

Dans le chapitre 5, nous proposons un modèle pour la classification **HR** à partir de l'information radiométrique temporelle décrite par une séquence d'images **BR** et de l'information spatiale fournie par une segmentation **HR**. Dans cet objectif, le modèle linéaire de mélange (cf. section 5.1) est utilisé pour décrire l'intensité mesurée à basse résolution en fonction de celle des différentes régions représentées dans le pixel basse résolution. L'utilisation du critère du *Maximum A Posteriori* (section 5.2) nous permet ensuite de définir une fonction d'énergie dont la minimisation conduit à l'élaboration d'une classification haute résolution. En faisant une hypothèse supplémentaire, cette énergie peut être simplifiée dans l'objectif de développer une méthode numériquement plus rapide. La stabilité de cette dernière énergie est étudiée dans la section 5.4.

La mise en œuvre de cette méthode et l'analyse de ses performances sont ensuite décrites dans le chapitre 6. Dans la section 6.1, nous décrivons l'algorithme de recuit simulé utilisé pour minimiser l'énergie considérée. La simulation des données et la démarche adoptée pour valider la méthode et générer une classification de référence sont ensuite précisées dans les sections 6.2 et 6.3. Enfin, nous présentons les premiers résultats obtenus sur des séquences d'images (section 6.4) avant d'analyser empiriquement la convergence de l'algorithme (section 6.5) ainsi que la sensibilité de la méthode vis-à-vis du rapport de résolution considéré (section 6.6).

## 5. Un modèle bayésien pour la classification

Dans ce chapitre, nous proposons un modèle bayésien pour la classification de l'occupation du sol à une échelle sous-pixellique, dans un contexte multi-temporel. Dans cet objectif, le modèle linéaire de mélange est utilisé pour décrire la séquence d'images basse résolution en termes du mélange des différentes régions de la segmentation contenues dans chaque pixel. La règle de Bayes et le critère du Maximum A Posteriori conduisent ensuite à la définition d'une fonction d'énergie dont le minimum correspond à la classification recherchée.

### 5.1. Modélisation du problème

La principale difficulté liée à l'utilisation d'images de résolution moyenne ou grossière telle que celles provenant des capteurs **MeRIS** ou **SPOT/VGT** vient de la taille du champ de vision d'un pixel relativement à la taille des objets d'intérêt. En effet, lorsqu'un pixel représente  $300 \times 300\text{m}$  ou  $1\text{km}^2$ , plusieurs parcelles sont alors susceptibles d'être représentées au sein de ce pixel. En présence de pixels mixtes (*i.e.* contenant plusieurs objets classes différentes, par opposition aux pixels purs), et sans information complémentaire, une séquence temporelle **BR** peut contraindre la solution (classification) uniquement en terme du pourcentage des classes représentées au sein de chaque pixel basse résolution. La localisation spatiale des objets à l'intérieur du pixel **BR** n'est alors pas accessible.

Dans le cas des zones agricoles, l'information spatiale correspond essentiellement à la structure géométrique de la scène (parcellaire, cadastre), qui peut être considérée comme relativement invariante au cours du temps. Nous supposons, dans cette partie, l'information spatiale constante au cours d'une année agricole et entièrement représentée sous la forme, par exemple, d'une segmentation ou d'un cadastre. En accord avec cette hypothèse, nous proposons d'aborder le problème de la classification en exploitant la richesse de l'information spatiale des images **HR** pour extraire les éléments géométriques (géographiques) d'une scène avant d'utiliser l'information temporelle des séquences d'images **BR** pour regrouper les différentes régions spatiales qui évoluent selon un même profil, caractéristique d'un type d'occupation du sol. Pour cela, nous décomposons le problème de la classification en deux problèmes successifs : la segmentation de la scène qui vise à en extraire les caractéristiques spatiales et sa labellisation.



Le problème de la segmentation a donné lieu à de nombreux travaux en analyse d'images, il n'est pas l'objet de cette étude. Par conséquent, nous considérons qu'une segmentation **HR** ou un cadastre (par exemple, de la base de données Corine land cover) est disponible. Dans le cas contraire, nous supposons qu'une image **HR**, au moins, est disponible et qu'une méthode de segmentation nous fournit des régions spatialement homogènes. Ici, nous utilisons une méthode de croissance des régions [Koepler et al., 1994] (module segct du logiciel <http://cm1a.ens-cachan.fr/Cm1a/Megawave>) permettant de fournir une partition de l'image avec un unique paramètre : le nombre de régions recherchées (ou la longueur des bords). Elle repose sur la fonctionnelle de [Mumford et Shah, 1989] et sur un algorithme multi-échelle et pyramidal qui semblent bien adaptés à un grand nombre d'images satellitaires.

Soit  $\mathcal{D}_{HR}$  le domaine haute résolution spatiale, nous considérons la définition 5.1.1.

**Définition 5.1.1 (Segmentation)** *Une segmentation du domaine  $\mathcal{D}_{HR}$  en  $S$  régions homogènes est une application  $s$ , constante par morceaux, définie par*

$$\begin{aligned} s : \mathcal{D}_{HR} &\rightarrow \mathcal{S} = \{1, \dots, S\} \\ x &\mapsto s(x). \end{aligned}$$

Avec cette partition de l'image en régions homogènes, notre objectif est alors d'exploiter l'information radiométrique temporelle de la séquence **BR** pour discriminer les types d'occupation du sol représentés dans chaque région. La définition 5.1.2 décrit le processus de labellisation.

**Définition 5.1.2 (Labellisation)** *On appelle labellisation  $\lambda$  une application qui associe un label (pris dans l'ensemble prédéfini des labels  $\mathcal{L}$ ) à toute région de la segmentation, soit*

$$\begin{aligned} \lambda : \mathcal{S} &\rightarrow \mathcal{L} = \{1, \dots, L\} \\ k &\mapsto \lambda_k. \end{aligned} \tag{5.1.1}$$

On note  $\lambda = (\lambda_1, \dots, \lambda_S)$  où, pour tout  $k \in \mathcal{S}$ ,  $\lambda_k$  désigne le label de la région  $k$ .

En pratique, la segmentation et la labellisation font aussi références, respectivement, à l'image des régions et des labels qui découlent de ces applications.

Soit  $\mathcal{T} = \{1, \dots, T\}$  l'ensemble des dates d'acquisition, la séquence d'images **BR** observée est notée sous la forme d'un vecteur d'images  $v = (v_1, \dots, v_T)$  où, pour tout  $t \in \mathcal{T}$ , l'image  $v_t$  est définie sur  $\mathcal{D}_{BR}$ . De manière analogue, une séquence d'images **HR** non-observée est notée  $u = (u_1, \dots, u_T)$  où, pour tout  $t \in \mathcal{T}$ , l'image  $u_t$  est définie sur  $\mathcal{D}_{HR}$ .

Dans un premier temps, nous nous intéressons à la description des relations entre une labellisation **HR** est une série d'observations **BR**. Pour cela, supposons que pour toute date  $t$ , une image **BR**  $v_t$  correspond à la moyenne par blocs d'une image **HR**,  $u_t$ , de la même scène. Pour tout  $y \in \mathcal{D}_{BR}$ , l'intensité mesurée à basse résolution vaut alors

$$v_t(y) = \frac{1}{N} \sum_{\{x \in y\}} u_t(x), \tag{5.1.2}$$

où  $\{x \in y\}$  représente l'ensemble des pixels haute résolution  $x \in \mathcal{D}_{\text{HR}}$  contenus dans le pixel basse résolution  $y$  et  $N = |\mathcal{D}_{\text{HR}}|/|\mathcal{D}_{\text{BR}}|$  est le nombre de pixels haute résolution contenus dans un pixel basse résolution ( $N$  est supposé entier). Si la labellisation est connue, remarquons que (5.1.2) peut alors être décomposée sous la forme

$$v_t(y) = \frac{1}{N} \sum_{l \in \mathcal{L}} \sum_{\substack{\{x \in y\} \\ \lambda_s(x) = l}} u_t(x), \quad (5.1.3)$$

où  $\lambda_s(x)$  représente le label du pixel  $x$ . La mesure basse résolution est alors décomposée par label. Soit  $N_l(y)$  le nombre de pixels haute résolution de label  $l$  représentés au sein du pixel  $y$ . La proportion du label  $l$  dans le pixel  $y$  vaut alors

$$\alpha_l(y) = \frac{N_l(y)}{N}, \quad (5.1.4)$$

avec la contrainte, pour tout  $y \in \mathcal{D}_{\text{BR}}$ , que

$$\sum_{l \in \mathcal{L}} \alpha_l(y) = 1. \quad (5.1.5)$$

Dans un contexte probabiliste, les observations  $u$  et  $v$  sont considérées comme des réalisations, respectivement, des champs aléatoires  $U$  et  $V$  où  $U = (U_1, \dots, U_T)$ ,  $V = (V_1, \dots, V_T)$ , et quelque soit  $t \in \mathcal{T}$ ,  $U_t = (U_t(x))_{x \in \mathcal{D}_{\text{HR}}}$  et  $V_t = (V_t(y))_{y \in \mathcal{D}_{\text{BR}}}$ . De plus, une labellisation  $\lambda = (\lambda_1, \dots, \lambda_S)$  est considérée en tant que réalisation du champ aléatoire des labels  $\Lambda = (\Lambda_1, \dots, \Lambda_S)$ . Avec ces notations et d'après (5.1.3), la variable aléatoire  $(V_t(y)|\Lambda = \lambda)$  vérifie alors, pour tout  $t \in \mathcal{T}$  et  $y \in \mathcal{D}_{\text{BR}}$ ,

$$\begin{aligned} (V_t(y)|\Lambda = \lambda) &= \sum_{l \in \mathcal{L}} \frac{N_l(y)}{N} \sum_{\substack{x \in \{x \in y\} \\ \lambda_s(x) = l}} \frac{(U_t(x)|\Lambda = \lambda)}{N_l(y)} \\ &= \sum_{l \in \mathcal{L}} \alpha_l(y) \sum_{\substack{\{x \in y\} \\ \lambda_s(x) = l}} \frac{(U_t(x)|\Lambda = \lambda)}{N_l(y)}. \end{aligned} \quad (5.1.6)$$

Dans la communauté de télédétection, l'équation (5.1.6) qui décrit la mesure observée en un pixel en fonction des différents types d'occupation du sol qui y sont représentés et de leur taux d'occupation est connue sous le nom de modèle linéaire de mélange. Ce modèle, introduit par [Horwitz et al., 1971] dans l'objectif initial d'estimer les proportions des différents objets représentés au sein d'un seul pixel, affirme que l'intensité observée en un pixel mixte peut être estimée par la moyenne des intensités qui auraient été mesurées dans un pixel pur représentant chaque type de couvert, pondérées par leur taux d'occupation dans le pixel. Plus précisément, pour toute date  $t \in \mathcal{T}$ , si la séquence d'images **BR**  $v$  représente une surface contenant au plus  $L$  types de couvert décrits par l'ensemble  $\mathcal{L}$ , le modèle linéaire de mélange permet d'exprimer l'intensité mesurée en tout pixel  $y \in \mathcal{D}_{\text{BR}}$  à partir des proportions et intensités caractéristiques de chaque composante sous la forme

$$v_t(y) = \sum_{l \in \mathcal{L}} \alpha_l(y) m_t(l), \quad (5.1.7)$$

où  $m_t(l)$  représente l'intensité caractéristique du type de couvert  $l$  à la date  $t$ . En pratique, l'intensité utilisée pour caractériser les différents types de couvert est souvent une intensité moyenne, et la relation obtenue par le modèle physique (5.1.7) revient à celle obtenue avec hypothèse de moyennage (5.1.6).

Dans la littérature, ce modèle est généralement appliqué à des images **BR**, dans un contexte où soit les proportions  $\alpha_l(y)$ , soit les intensités caractéristiques  $m_t(l)$  sont inconnues. Certains travaux l'utilisent pour estimer les proportions connaissant un *a priori* ([Horwitz et al., 1971, Settle et Drake, 1993]) et d'autres pour estimer les intensités caractéristiques de chaque type de couvert connaissant les proportions ([Cross et al., 1991]). D'autres travaux, plus récents, considèrent alternativement les proportions  $\alpha$  et les intensités caractéristiques  $m$  comme inconnues, de manière à estimer ces deux paramètres par un algorithme de type EM ou ICE (*cf.* section 3.3).

Dans le cas général, les valeurs des images **BR** sont obtenues par une Fonction de Transfert de Modulation (FTM) qui ne correspond pas à une fonction "porte" dans le domaine spatial comme le suppose l'hypothèse de moyennage. Cependant, un certain nombre de travaux expérimentaux montrent, en pratique, que le modèle linéaire de mélange est raisonnable pour de nombreuses applications. Notamment, [Kerdiles et Grondona, 1995, Small, 2001] ont montré que l'utilisation de la combinaison linéaire des valeurs de NDVI, bien que non linéaires, engendrait des erreurs mineures. Par la suite, le modèle linéaire de mélange fera référence à cette relation, qu'elle provienne de l'hypothèse de moyennage ou du modèle physique. Cependant, une approximation plus fine de la FTM d'un capteur donné serait possible par un modèle linéaire défini en prenant en compte les recouvrements de spectre.

En considérant que la valeur radiométrique mesurée à un instant donné dépend des réponses spectrales des différents types d'occupation du sol représentés au sein du champ de vision, le problème de la labellisation peut être envisagé, dans un cadre probabiliste, sous la forme d'un problème d'estimation d'une réalisation non-observée du champ aléatoire des labels à partir de l'observation de données bruitées. Dans cet objectif, nous nous plaçons dans un cadre bayésien et nous considérons le critère du *Maximum A Posteriori* (MAP). Cet estimateur repose sur une fonction de coût de type 0 – 1, *i.e.* pour laquelle toutes les configurations sont pénalisées de la même manière sauf la labellisation correcte. Avec cette fonction de coût, la labellisation qui minimise le coût moyen est alors celle qui maximise la probabilité *a posteriori* d'obtenir une certaine configuration de labels connaissant la série temporelle observée. Cette probabilité *a posteriori* peut être exprimée, avec la règle de Bayes, en fonction de la probabilité d'observer une séquence d'images étant donnée une labellisation et de la probabilité *a priori* d'obtenir une certaine labellisation.

Dans la section 5.2, une hypothèse gaussienne sur la distribution des observations conditionnellement à une labellisation permet de déterminer la probabilité conditionnelle *a priori*. Le calcul de la probabilité *a posteriori* d'obtenir une labellisation étant donnée une observation permet ensuite de déterminer la labellisation optimale au sens du MAP.

## 5.2. Maximum *a posteriori*

Dans cette section, notre objectif est de déterminer la probabilité *a posteriori* d'obtenir une certaine labellisation étant donnée une séquence d'images observée à basse résolution et une segmentation haute résolution. La maximisation de cette probabilité permettra alors de générer une labellisation sous-pixellique (à haute résolution spatiale).

Avec le modèle linéaire de mélange (cf. Section 5.1), la mesure observée au sein d'un pixel basse résolution est exprimée en fonction du taux d'occupation de chaque label dans le pixel d'intérêt et des moyennes caractéristiques des différents labels. Dans l'objectif d'attribuer un label à chaque région de la segmentation, considérons  $\beta_k(y)$  la proportion de la région  $k$  représentée par le pixel  $y$ . Pour tout label  $l \in \mathcal{L}$ , la proportion du label  $l$  dans le pixel  $y$  est alors égale à la somme des proportions de toutes les régions de label  $l$  présentées dans le pixel  $y$ , soit

$$\alpha_l(y) = \sum_{\substack{k \in S \\ \lambda_k = l}} \beta_k(y), \quad (5.2.8)$$

où  $S$  représente le nombre de régions existantes dans la segmentation et  $\lambda_k$  le label de la région  $k$ . Nous supposons que le nombre  $L$  des différents types d'occupation du sol présents sur la surface géographique étudiée est connu.

La probabilité d'observer  $v$  sachant la labellisation  $\lambda$  décrit les processus d'observation et d'acquisition des données. Pour la déterminer, nous supposons que, pour tout  $t$ , le champ aléatoire  $U_t$  sur le domaine haute résolution  $\mathcal{D}_{HR}$  est gaussien conditionnellement à la labellisation haute résolution. Pour toute date  $t$ , le type d'occupation du sol de label  $l$  est alors caractérisé par une distribution gaussienne de moyenne  $m_t(l)$  et de variance  $var_t(l)$ . La variance  $var_t(l)$  modélise la variabilité intrinsèque d'une classe donnée. Cette hypothèse est couramment utilisée dans le cas des données de télédétection optiques. Les mesures acquises par les capteurs haute et basse résolution étant liées par un opérateur de moyennage (i.e. l'image  $V_t$  correspond à la moyenne par blocs de l'image  $U_t$ ), la variable aléatoire  $V_t(y)$  est aussi gaussienne conditionnellement à la labellisation.

En général, les images acquises à basse résolution comportent de nombreux pixels mixtes, la taille des objets d'intérêt étant souvent plus petite que la surface représentée dans le champ de vision d'un pixel. Il est alors difficile de faire une hypothèse de régularité spatiale entre pixels **BR** voisins. Par conséquent, nous supposons que les observations **BR** sont spatialement indépendantes conditionnellement à la labellisation. Les observations **HR**, quant à elles, ne sont pas considérées directement mais uniquement à travers les mesures **BR**. Elles sont donc à leur tour supposées indépendantes conditionnellement à la labellisation. La distribution de la loi *a priori* est alors déterminée par la proposition 5.2.1.

**Proposition 5.2.1** Soient  $s$  une segmentation et  $(\beta_k)_{k \in S}$  les images de proportions **BR** associées. Supposons que, pour toute date  $t \in \mathcal{T}$  et tout pixel  $x \in \mathcal{D}_{HR}$ , la variable aléatoire  $U_t(x)$  suit, conditionnellement à une labellisation  $\Lambda = \lambda$ , une loi normale de moyenne  $m_t(\lambda_{s(x)})$  et de variance

$var_t(\lambda_{s(x)})$ . Sous hypothèse d'indépendance des variables  $(U_t(x))_{x \in \mathcal{D}_{HR}}$ , la densité a priori de  $V_t(y)$  conditionnellement à  $\Lambda = \lambda$  vaut alors

$$f_{V_t(y)|\Lambda}(v_t(y)|\lambda) = \frac{1}{\sigma_t(\lambda, y)\sqrt{2\pi}} \exp\left(-\frac{(v_t(y) - \mu_t(\lambda, y))^2}{2\sigma_t^2(\lambda, y)}\right), \quad (5.2.9)$$

densité gaussienne de moyenne

$$\mu_t(\lambda, y) = \sum_{l \in \mathcal{L}} \sum_{\substack{k \in \mathcal{S} \\ l_k = l}} \beta_k(y) m_t(l), \quad (5.2.10)$$

et de variance

$$\sigma_t^2(\lambda, y) = \frac{1}{N} \sum_{l \in \mathcal{L}} \sum_{\substack{k \in \mathcal{S} \\ l_k = l}} \beta_k(y) var_t(l). \quad (5.2.11)$$

### Démonstration.

En supposant que, pour tout  $t \in \mathcal{T}$  et  $x \in \mathcal{D}_{HR}$ , la variable aléatoire  $(U_t(x)|\Lambda = \lambda)$  suit une loi normale de moyenne  $m_t(\lambda_{s(x)})$  et de variance  $var_t(\lambda_{s(x)})$ , alors le modèle linéaire de mélange (5.1.7) permet d'exprimer l'intensité observée sous la forme :

$$(V_t(y)|\Lambda = \lambda) = \sum_{l \in \mathcal{L}} \alpha_l(y) \sum_{\substack{\{x \in y\} \\ \lambda_{s(x)} = l}} \frac{(U_t(x)|\Lambda = \lambda)}{N_l(y)}, \quad (5.2.12)$$

où  $N_l(y)$  représente le nombre de pixels  $x \in \mathcal{D}_{HR}$  de label  $l$  présents dans le pixel basse résolution  $y$  et  $\{x \in y\}$  désigne l'ensemble des pixels haute résolution représentés au sein du pixel  $y$ . La variable  $(V_t(y)|\Lambda = \lambda)$  est donc gaussienne en tant que mélange de variables gaussiennes. Sa moyenne vaut

$$\begin{aligned} \mathbb{E}[(V_t(y)|\Lambda = \lambda)] &= \mathbb{E}\left[\sum_{l \in \mathcal{L}} \alpha_l(y) \sum_{\substack{\{x \in y\} \\ \lambda_{s(x)} = l}} \frac{(U_t(x)|\Lambda = \lambda)}{N_l(y)}\right] \\ &= \sum_{l \in \mathcal{L}} \alpha_l(y) \sum_{\substack{\{x \in y\} \\ \lambda_{s(x)} = l}} \frac{\mathbb{E}[(U_t(x)|\Lambda = \lambda)]}{N_l(y)} \\ &= \sum_{l \in \mathcal{L}} \alpha_l(y) \sum_{\substack{\{x \in y\} \\ \lambda_{s(x)} = l}} \frac{m_t(l)}{N_l(y)} \\ &= \sum_{l \in \mathcal{L}} \alpha_l(y) m_t(l) \text{ car } N_l(y) = |\{\{x \in y\}, \lambda_{s(x)} = l\}| \\ &= \sum_{l \in \mathcal{L}} \sum_{\substack{k \in \mathcal{S} \\ l_k = l}} \beta_k(y) m_t(l), \text{ (par (5.2.8)).} \end{aligned} \quad (5.2.13)$$

De même, la variance est calculée à partir de l'équation (5.1.6) et, grâce à l'hypothèse d'indépendance des variables  $(U_t(x))_{x \in \mathcal{D}_{\text{HR}}}$ , elle vaut, pour toute date  $t$  et tout pixel  $y$ ,

$$\begin{aligned}
 \sigma_t^2(\lambda, y) &= \sum_{l \in \mathcal{L}} \alpha_l^2(y) N_l(y) \left( \frac{1}{N_l(y)} \right)^2 \text{var}_t(l) \\
 &= \frac{1}{N} \sum_{l \in \mathcal{L}} \alpha_l(y) \text{var}_t(l) \\
 &= \frac{1}{N} \sum_{l \in \mathcal{L}} \sum_{\substack{k \in S \\ l_k = l}} \beta_k(y) \text{var}_t(l).
 \end{aligned} \tag{5.2.14}$$

□

Remarquons que les moyenne et variance dépendent alors du pixel basse résolution considéré, à travers sa composition en terme des différents types d'occupation du sol. La dépendance à la labellisation se manifeste alors uniquement par l'expression du mélange des labels au sein des pixels basse résolution.

Supposons, de plus, l'indépendance temporelle des observations conditionnellement à la labellisation. Si cette hypothèse semble peu réaliste, elle est courante dans la littérature dans la mesure où l'estimation des corrélations croisées entre les différentes dates considérées est difficile à réaliser et nécessiterait une phase d'apprentissage ou la prise en compte d'un *a priori* sur le profil temporel des différents types d'occupation du sol. D'un point de vue théorique la prise en compte des corrélations temporelles n'engendre pas de difficulté particulière. Elle se manifeste simplement sous la forme d'une matrice de covariance non diagonale. Cependant, l'estimation des termes de covariance temporelle risque alors de souffrir du phénomène de Hughes selon lequel, pour un nombre limité d'échantillons, la qualité de l'estimation décroît lorsque la dimension augmente. Dans notre cas, le nombre de termes de covariance à estimer devient rapidement très grand relativement à la taille de l'échantillon constitué par les images **BR**. Si l'estimation d'un plus grand nombre de paramètres (les termes non-diagonaux) permet de diminuer le biais de l'estimateur, elle en augmente la variance (*i.e.* l'erreur stochastique). Il est donc nécessaire de faire un compromis entre ces deux types d'erreurs.

Sous ces hypothèses, la densité de probabilité d'observer une mesure  $v$  est déterminée, conditionnellement à la labellisation, par

$$f_{V|\Lambda=\lambda}(v|\lambda) = \prod_{t \in T} \prod_{y \in \mathcal{D}_{\text{BR}}} f_{V_t(y)|\Lambda=\lambda}(v_t(y)|\lambda). \tag{5.2.15}$$

La moyenne du champ aléatoire  $V$  conditionnellement à la labellisation est alors, pour tout pixel  $y$ , notée  $\mu(\lambda, y)$ . Elle correspond au vecteur des moyennes  $(\mu_t(\lambda, y))_{y,t}$ . Quant à la matrice de covariance, elle est diagonale grâce à l'hypothèse d'indépendance temporelle et elle peut s'écrire sous la forme  $\sigma^2(\lambda, y) I_T$ , où  $\sigma^2(\lambda, y) = (\sigma_t^2(\lambda, y))_{y,t}$  et  $I_T$  désigne la matrice identité en dimension  $T$ .

Nous proposons de considérer le critère du *Maximum A Posteriori* pour estimer la configuration optimale du champ des labels  $\Lambda$  connaissant la séquence d'observation  $v$ . Notons  $\lambda^* = (\lambda_1^*, \dots, \lambda_S^*)$  la labellisation optimale, où la  $k$ ième coordonnée représente le label de la  $k$ ième région.

En l'absence d'information contextuelle liée à la scène observée, nous supposons que tous les labels ont la même probabilité d'occurrence. En pratique, cette hypothèse est peu réaliste. Par exemple, un label représentant une classe telle que l'eau est généralement très minoritaire dans une scène rurale. Dans la littérature, un formalisme markovien pour le champ des labels est souvent adopté pour modéliser les interactions entre pixels voisins sous la forme d'énergies locales dans l'objectif de favoriser l'exploration de certaines solutions qui correspondent à un *a priori* sur l'image des labels (cf. [Geman et Geman, 1984, Kasetkasem et al., 2005]). Dans notre cas, une hypothèse de régularité spatiale est difficile à envisager puisque les pixels basse résolution contiennent *a priori* plus d'un objet d'intérêt (pixels mixtes). Cependant, l'utilisation d'une segmentation **HR** permet de contraindre l'exploration des champs de labels.

Ces hypothèses permettent d'obtenir, avec la proposition 5.2.2, l'expression de la labellisation optimale en fonction de l'observation, des moyennes et variances de chaque label et des taux d'occupation de chaque région au sein de chaque pixel.

**Proposition 5.2.2** *Sous les hypothèses de normalité et d'indépendance de la proposition 5.2.1, la labellisation optimale au sens du Maximum A Posteriori est déterminée par*

$$\lambda^* = \arg \min_{\lambda \in L^S} \sum_{t \in T} \sum_{y \in \mathcal{D}_{BR}} \left( \frac{(v_t(y) - \mu_t(\lambda, y))^2}{\sigma_t^2(\lambda, y)} + \ln(\sigma_t^2(\lambda, y)) \right) \quad (5.2.16)$$

**Démonstration.**

D'après la règle de Bayes, la densité de probabilité *a posteriori* s'écrit

$$f_{\Lambda|V=v}(\lambda|v) = \frac{f_{V|\Lambda}(v|\lambda) f_{\Lambda}(\lambda)}{f_V(v)}. \quad (5.2.17)$$

La densité  $f_V$  est une constante pour le problème de maximisation considéré. De plus, l'hypothèse d'équidistribution des labels dans l'image implique que  $f_{\Lambda}$  ne joue alors pas de rôle sur la maximisation. Par conséquent,

$$\arg \max_{\lambda \in L^S} f_{\Lambda|V=v}(\lambda|v) = \arg \max_{\lambda \in L^S} f_{V|\Lambda=\lambda}(v|\lambda) \quad (5.2.18)$$

et la labellisation optimale est alors déterminée par

$$\begin{aligned} \lambda^* &= \arg \max_{\lambda \in L^S} f_{V|\Lambda=\lambda}(v|\lambda) \\ &= \arg \max_{\lambda \in L^S} \prod_{t \in T} \prod_{y \in \mathcal{D}_{BR}} f_{V_t(y)|\Lambda=\lambda}(v_t(y)|\lambda) \\ &= \arg \max_{\lambda \in L^S} \prod_{t \in T} \prod_{y \in \mathcal{D}_{BR}} \frac{1}{\sigma_t(y, l) \sqrt{2\pi}} \exp \left( -\frac{(v_t(y) - \mu_t(\lambda, y))^2}{2\sigma_t^2(y, l)} \right) \\ &= \arg \min_{\lambda \in L^S} \sum_{t \in T} \sum_{y \in \mathcal{D}_{BR}} \left( \frac{(v_t(y) - \mu_t(\lambda, y))^2}{\sigma_t^2(\lambda, y)} + \ln(\sigma_t^2(\lambda, y)) \right). \end{aligned}$$

□

Avec l'hypothèse d'équidistribution des labels dans l'image, remarquons que le critère du MAP est équivalent au critère du maximum de vraisemblance. La résolution du problème (5.2.16) conduit alors à une labellisation optimale au sens du Maximum *A Posteriori* ou du maximum de vraisemblance. Elle est, cependant, conditionnée par la connaissance des moyennes et variances caractéristiques des différentes classes d'occupation du sol présentes sur la surface considérée. En pratique, si les moyennes et variances caractéristiques des classes sont connues *a priori*, une solution de (5.2.16) peut être obtenue par un processus d'optimisation global. On parle alors de méthode de classification supervisée. Lorsqu'elles sont inconnues, elles peuvent être estimées au fil du processus d'optimisation grâce au modèle linéaire de mélange, au prix d'un temps de calcul plus important.

En général, les caractéristiques des classes sont inconnues : elles dépendent des dates d'acquisition, des conditions atmosphériques, de la croissance de la végétation à une échelle locale (le cycle phénologique d'une même culture n'est pas nécessairement identique sur deux sites géographiques distincts), etc. Une méthode de classification non-supervisée semble alors plus appropriée. Les caractéristiques des classes peuvent alors être estimées à partir du modèle linéaire de mélange, connaissant les proportions des différentes classes au sein de chaque pixel. Cependant, si les moyennes des classes peuvent être estimées en un temps raisonnable lors du processus de minimisation, par exemple au moyen d'une régression linéaire, l'estimation des variances de classes est plus sensible et coûteuse en temps de calcul.

Par conséquent, pour une approche non-supervisée, nous proposons de nous affranchir du problème de l'estimation des variances des classes et de sa robustesse en supposant une variance,  $\sigma^2$ , égale pour toutes les classes. Cette hypothèse conduit à une variance constante pour tout pixel basse résolution. Le corollaire 5.2.3 permet de déterminer la labellisation optimale correspondante.

**Corollaire 5.2.3** *Sous les hypothèses précédentes, la labellisation optimale au sens du Maximum *A Posteriori* est déterminée par*

$$\lambda^* = \arg \min_{\lambda \in \mathcal{L}^S} \sum_{t \in T} \sum_{y \in \mathcal{D}_{BR}} (v_t(y) - \mu_t(\lambda, y))^2. \quad (5.2.19)$$

#### Démonstration.

Si les variances de tous les labels sont égales, alors la variance d'un pixel basse résolution vaut

$$\begin{aligned} \sigma_t^2(\lambda, y) &= \frac{1}{N} \sum_{l \in \mathcal{L}} \alpha_l(y) \sigma^2 \\ &= \frac{\sigma^2}{N} \sum_{l \in \mathcal{L}} \alpha_l(y) \\ &= \frac{\sigma^2}{N}. \end{aligned}$$



D'après la proposition 5.2.2, la labellisation optimale vaut alors

$$\begin{aligned}\lambda^* &= \arg \min_{\lambda \in L^S} \sum_{t \in \mathcal{T}} \sum_{y \in \mathcal{D}_{\text{BR}}} \left( \frac{N}{\sigma^2} (v_t(y) - \mu_t(\lambda, y))^2 + \ln\left(\frac{\sigma^2}{N}\right) \right) \\ &= \arg \min_{\lambda \in L^S} \sum_{t \in \mathcal{T}} \sum_{y \in \mathcal{D}_{\text{BR}}} (v_t(y) - \mu_t(\lambda, y))^2.\end{aligned}$$

□

Par la suite, on note

$$E_v(\lambda) = \sum_{t \in \mathcal{T}} \sum_{y \in \mathcal{D}_{\text{BR}}} (v_t(y) - \mu_t(\lambda, y))^2, \quad (5.2.20)$$

l'énergie associée à une labellisation  $\lambda$  et une séquence d'images  $v$  considérée dans le problème (5.2.19). D'après quelques statistiques effectuées sur des données réelles (**SPOT/HRV**), cette hypothèse ne semble pas aberrante dans la mesure où les différentes classes sont généralement de variance proche à l'exception de la classe 'eau' dont la variance est très faible. Cependant, la classe 'eau' peut être nettement distinguée des autres classes par sa radiométrie et une sur-estimation de la variance n'est alors pas nuisible à la labellisation.

Dans la section 5.3, nous étudions les solutions du problème (5.2.19) et nous nous intéressons à l'influence de la basse résolution sur ces solutions.

### 5.3. Influence de la basse résolution

Dans cette section, notre objectif est d'évaluer la perte d'information résultant de l'utilisation de données basse résolution par rapport à l'utilisation de données haute résolution en considérant le problème (5.2.19).

Dans un contexte non-supervisé, les moyennes caractéristiques des classes sont inconnues. Le problème (5.2.19) peut alors être utilisé conjointement pour l'estimation des moyennes et de la labellisation. Pour simplifier les écritures, nous adoptons ici les notations matricielles suivantes :

- $A \in \mathbb{R}^{|\mathcal{D}_{\text{HR}}|}$  le vecteur des données **HR** ( $u(x)_{x \in \mathcal{D}_{\text{HR}}}$ ),
- $B \in \mathcal{M}(|\mathcal{D}_{\text{BR}}|, |\mathcal{D}_{\text{HR}}|)$  l'opérateur de moyennage par blocs reliant les images **HR** et **BR**,
- $R \in \mathcal{M}(|\mathcal{D}_{\text{HR}}|, S)$  la matrice qui correspond à la segmentation : le terme en  $(i, j)$  vaut 1 si le pixel  $i$  appartient à la région  $j$  et 0 sinon,
- $C \in \mathcal{M}(S, L)$  la matrice qui correspond à la labellisation : le terme en  $(i, j)$  vaut 1 si la région  $i$  est de label  $j$ , 0 sinon,
- $M \in \mathbb{R}^L$  le vecteur des moyennes des classes.

Avec ces nouvelles notations, le problème considéré par (5.2.19) s'écrit sous la forme

$$\inf_{C, M} \|B(RCM - A)\|^2, \quad (5.3.21)$$

et le problème analogue sans hypothèse de mélange est défini par

$$\inf_{C, M} \|RCM - A\|^2. \quad (5.3.22)$$

Remarquons que pour une labellisation fixée, les problèmes des moindres carrés (5.3.21) et (5.3.22) admettent chacun au moins une solution.

**Proposition 5.3.1** *Soit  $C \in \mathcal{M}(S, L)$  une labellisation fixée.*

– *Il existe au moins un vecteur  $M \in \mathbb{R}^L$  tel que*

$$\|RCM - A\| = \inf_{X \in \mathbb{R}^L} \|RCX - A\|. \quad (5.3.23)$$

*Cette solution est unique si et seulement si l'application qui correspond à la matrice  $RC$  est injective.*

– *Il existe au moins un vecteur  $M \in \mathbb{R}^{|\mathcal{D}_{HR}|}$  tel que*

$$\|B(RCM - A)\| = \inf_{X \in \mathbb{R}^L} \|B(RCX - A)\|. \quad (5.3.24)$$

*Cette solution est unique si et seulement si l'application qui correspond à la matrice  $BRC$  est injective.*

**Démonstration.**

Les deux points de la proposition découlent directement du théorème de projection (cf. [Ciarlet, 1982]).

**Théorème 5.3.2 (théorème de projection)** *Soit  $U$  un sous-ensemble non vide, convexe, fermé, d'un espace de Hilbert  $V$ . Etant donné un élément quelconque  $w \in V$ , il existe un et un seul élément  $Pw$  tel que*

$$Pw \in U \text{ et } \|w - Pw\| = \inf_{u \in U} \|w - u\|. \quad (5.3.25)$$

Considérons le problème

$$\inf_{X \in \mathbb{R}^L} \|RCX - A\|^2, \quad (5.3.26)$$

où la matrice  $RC \in \mathcal{M}(|\mathcal{D}_{HR}|, L)$  et le vecteur  $A \in \mathbb{R}^{|\mathcal{D}_{HR}|}$  sont donnés. Le sous-espace vectoriel  $\text{Im}(RC) = \{RCX \in \mathbb{R}^{|\mathcal{D}_{HR}|}; X \in \mathbb{R}^L\}$  est fermé. Le théorème de projection entraîne alors l'existence et l'unicité d'un élément  $Y$  tel que  $Y \in \text{Im}(RC)$  et  $\|Y - A\| = \inf_{Z \in \text{Im}(RC)} \|Z - A\|$ . Le problème (5.3.23) admet donc au moins une solution : un vecteur  $M \in \mathbb{R}^L$  tel que  $RCM = Y$ . L'ensemble des solutions est alors défini par  $\{M + e, e \in \text{Ker}(RC)\}$ . Le problème (5.3.23) admet donc une solution unique si et seulement si  $\text{Ker}(RC) = 0$ , c'est à dire si et seulement si l'application représentée par la matrice  $RC$  est injective. De même pour le problème (5.3.24) en considérant la matrice  $BRC$  et le vecteur  $BA$ .  $\square$

Dans notre cas, la matrice  $RC$  correspond à l'application qui permet de reconstruire une image **HR** où la valeur en chaque pixel est la valeur moyenne associée au label de ce pixel. Par conséquent, tout

élément de  $\text{Im}(RC)$  a au plus un antécédent par l'application  $RC$  et, en général, l'application que nous considérons est injective. Dans le contexte que nous nous sommes fixé, les problèmes (5.3.23) et (5.3.24) admettent donc une solution unique.

La comparaison des vecteurs de moyennes solutions des problèmes (5.3.24) et (5.3.23) nous permet d'évaluer, dans un cadre théorique, la perte d'information due à l'utilisation de données basse résolution plutôt que de données haute résolution.

**Proposition 5.3.3** *Soit  $R \in \mathcal{M}(|\mathcal{D}_{HR}|, S)$  la matrice associée à la segmentation et  $C \in \mathcal{M}(S, L)$  la matrice associée à une labellisation fixée. Supposons que les matrices  $RC$  et  $BRC$  soient de rang  $L$ . Si les vecteurs  $M, M' \in \mathbb{R}^L$  sont solutions, respectivement, des problèmes des moindres carrés (5.3.23) et (5.3.24), ils vérifient*

$$\|M - M'\|^2 \leq \|(I - {}^tBB)A\|^2 \sum_{i=1}^L \frac{1}{d_i^2} \quad (5.3.27)$$

et

$$\|M - M'\|^2 \leq \left(1 - \frac{1}{N}\right)^2 \|A\|^2 \sum_{i=1}^L \frac{1}{d_i^2} \quad (5.3.28)$$

où  $\|\cdot\|$  est la norme matricielle subordonnée à la norme vectorielle  $\|\cdot\|$  et  $(d_i)_{i=1 \dots L}$  représente l'ensemble des valeurs singulières de la matrice  $RC$ .

#### Démonstration.

La décomposition en valeurs singulières de la matrice  $RC$  permet de l'exprimer sous la forme

$$RC = O D {}^tW. \quad (5.3.29)$$

où  $O$  est une matrice orthogonale de taille  $|\mathcal{D}_{HR}| \times L$ ,  $W$  est une matrice orthogonale de taille  $L \times L$  et  $D = (d_i)_i$  est une matrice diagonale de taille  $L \times L$  contenant les valeurs singulières de  $RC$  sur la diagonale. Comme le rang de  $RC$  vaut  $L$ , les valeurs singulières de  $RC$  sont toutes non-nulles. Notons  $O = (O_1, \dots, O_L)$  et  $W = (W_1, \dots, W_L)$  où pour  $i = 1 \dots L$ , les termes  $O_i$  et  $W_i$  désignent, respectivement, les vecteurs colonnes des matrices  $O$  et  $W$ . Pour tout  $X \in \mathbb{R}^L$ , on a alors

$$RCX = \sum_{i=1}^L d_i \langle X, W_i \rangle O_i. \quad (5.3.30)$$

Par conséquent, la solution de (5.3.23) (unique d'après la proposition 5.3.1) est définie par

$$M = \sum_{i=1}^L \frac{1}{d_i} \langle A, O_i \rangle W_i \quad (5.3.31)$$

et la moyenne  $M'$ , solution du problème (5.3.24), s'écrit alors

$$\begin{aligned} M' &= \sum_{i=1}^L \frac{1}{d_i} \langle BA, BO_i \rangle W_i \\ &= \sum_{i=1}^L \frac{1}{d_i} \langle {}^tBBA, O_i \rangle W_i. \end{aligned} \quad (5.3.32)$$

Donc

$$\begin{aligned}
 \|M - M'\|^2 &= \left\| \sum_{i=1}^L \frac{1}{d_i} \langle A - {}^tBBA, O_i \rangle W_i \right\|^2 \\
 &\leq \sum_{i=1}^L \frac{1}{d_i^2} \langle A - {}^tBBA, O_i \rangle^2 \\
 &\leq \sum_{i=1}^L \frac{1}{d_i^2} \|A(I - {}^tBB)\|^2
 \end{aligned} \tag{5.3.33}$$

$$\leq \sum_{i=1}^L \frac{1}{d_i^2} \|I - {}^tBB\|^2 \|A\|^2. \tag{5.3.34}$$

$$\tag{5.3.35}$$

Remarquons que la matrice

$$B = \begin{pmatrix} \frac{1}{N} \cdots \frac{1}{N} & 0 & \cdots & 0 \\ 0 & \frac{1}{N} \cdots \frac{1}{N} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{N} \cdots \frac{1}{N} \end{pmatrix}$$

où  $N = \frac{|\mathcal{D}_{\text{HR}}|}{|\mathcal{D}_{\text{BR}}|}$  et la valeur  $\frac{1}{N}$  apparaît  $N$  fois consécutives sur chaque ligne. La matrice  ${}^tBB$ , de

taille  $|\mathcal{D}_{\text{HR}}| \times |\mathcal{D}_{\text{HR}}|$ , vaut alors  $\frac{1}{N^2} J_N$  où  $J_N = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$ . Par conséquent,

$$\|I - {}^tBB\| = 1 - \frac{1}{N^2} + \frac{N-1}{N^2} = 1 - \frac{1}{N}$$

et

$$\|M - M'\|^2 \leq \left(1 - \frac{1}{N}\right)^2 \|A\|^2 \sum_{i=1}^L \frac{1}{d_i^2}. \tag{5.3.36}$$

□

D'après l'équation (5.3.32), remarquons le vecteur des moyennes obtenu à partir d'une observation **BR** correspond au vecteur qu'on aurait obtenu en observant  ${}^tBBA$ . L'image  ${}^tBBA$  représente l'image **HR** moyennée par blocs de taille  $N \times N$  (où  $N = \frac{|\mathcal{D}_{\text{HR}}|}{|\mathcal{D}_{\text{BR}}|}$ ). Pour une labellisation fixée, cette écriture permet de remarquer que si l'image **BR** zoomée est égale à l'image **HR**, alors les problèmes (5.3.23) et (5.3.24) admettent la même solution. Cette situation se produit, en particulier, dans le cas où tous les pixels basse résolution sont *purs*.

De manière plus générale, l'inégalité (5.3.27) permet d'évaluer le gain apporté par l'utilisation de données **BR**. En effet, le terme  $\|(I - {}^tBB)A\|^2$  correspond à l'erreur quadratique réalisée en observant  ${}^tBBA$  plutôt que  $A$ . Il peut être calculé pour une observation et un facteur

de résolution donnés. Lorsque les matrices  $R$  et  $C$  sont fixées, le terme  $\sum_{i=1}^L \frac{1}{d_i^2}$  peut aussi être déterminé, chacune des valeurs propres  $d_i$  de la matrice  ${}^t(RC)(RC)$  étant égale au nombre total de pixels de label  $i$ . Pour un jeu de données fixé, une majoration de l'erreur quadratique entre les vecteurs  $M$  et  $M'$  peut donc être obtenue, permettant ainsi d'évaluer l'erreur introduite par la prise en compte de données **BR**.

L'inégalité (5.3.28) permet d'évaluer l'erreur engendrée par l'exploitation de données **BR** en fonction du rapport de résolution  $N$  considéré. En effet, pour une observation **HR** donnée, le terme  $\|A\|^2$  peut être calculé directement. Lorsque les matrices  $R$  et  $C$  sont fixées, le terme  $\|A\|^2 \sum_{i=1}^L$  est alors constant et seul le terme  $(1 - \frac{1}{N})^2$  varie en fonction du rapport de résolution  $N$ . Il est compris entre 0 et 1 et croît avec le rapport de résolution.

Ces résultats permettent de comprendre l'influence de la basse résolution sur le problème de l'estimation des moyennes par label et d'évaluer le poids de cette influence en fonction du facteur de résolution considéré même si, en pratique, les problèmes que nous considérons sont plus compliqués puisque les matrices  $C$  et  $M$  sont toutes deux inconnues. Il s'agit alors de rechercher les matrices  $CM$  solutions des problèmes (5.3.21) et (5.3.22). Remarquons qu'une matrice  $CM$  est de taille  $S \times L$  et ses coefficients admettent au plus  $L$  valeurs. L'espace de ces solutions est très combinatoire et son étude est donc difficile.

Dans la section 5.4, nous nous intéressons à la probabilité d'obtenir une erreur de labellisation en résolvant le problème (5.3.21). Cette analyse permet, en particulier, de considérer les aspects multitemporels et leur influence sur cette probabilité.

## 5.4. Analyse d'erreurs et aspects multitemporels

Avant de nous intéresser à la mise en oeuvre du modèle de labellisation, nous proposons d'analyser l'apparition d'erreurs de labellisation en résolvant le problème (5.2.19). Pour cela, nous calculons la probabilité, sous les hypothèses du corollaire 5.2.3, d'attribuer un label incorrect à une région de la segmentation. Rappelons que  $v(y) = (v_t(y))_{t \in \mathcal{T}}$  désigne le vecteur des observations **BR** en un pixel  $y$ , réalisation du processus aléatoire  $V(y) = (V_t(y))_{t \in \mathcal{T}}$  et  $m(\lambda_k) = (m_t(\lambda_k))_{t \in \mathcal{T}}$  le vecteur temporel des moyennes associées au label de la région  $k$ . Pour calculer la probabilité de faire une erreur de labellisation, considérons l'énergie (5.2.20) du problème (5.2.19) associée à la variable aléatoire  $V$  :

$$E_V(\lambda) = \sum_{y \in \mathcal{D}_{\text{BR}}} \|V(y) - \sum_{k \in \mathcal{S}} \beta_k(y) m(\lambda_k)\|^2, \quad (5.4.37)$$

où  $\|\cdot\|$  désigne la norme euclidienne.

**Définition 5.4.1** Soit  $\lambda = (\lambda_1, \dots, \lambda_S)$  la labellisation correcte. Une labellisation  $\lambda' = (\lambda'_1, \dots, \lambda'_S)$  est dite 1-optimale si elle vérifie

- .  $\forall k \neq k_0, \lambda'_k = \lambda_k$  ,
- .  $\lambda'_{k_0} \neq \lambda_{k_0}$  ,

où la région  $k_0$  est entièrement contenue dans un pixel  $\mathbf{BR} y_0$ . On note  $\beta_{k_0}(y_0)$  le taux d'occupation de la région  $k_0$  dans le pixel  $y_0$ .

La proposition 5.4.2 détermine la probabilité qu'une labellisation 1-optimale soit préférée à la labellisation correcte.

**Proposition 5.4.2** *Soit  $V$  un champ aléatoire gaussien vérifiant les hypothèses du corollaire 5.2.3. La probabilité pour qu'une labellisation 1-optimale  $\lambda'$  soit préférée à la labellisation optimale  $\lambda$  satisfait*

$$\mathbb{P}((E(\lambda) - E(\lambda')) \geq 0) = G\left(-\frac{\sqrt{N}\|\delta(y_0)\|}{2\sigma}\right), \quad (5.4.38)$$

où

$$\|\delta(y_0)\| = \beta_{k_0}(y_0) \|m(\lambda_{k_0}) - m(\lambda'_{k_0})\|, \quad (5.4.39)$$

et

$$\forall x \in \mathbb{R}, G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \quad (5.4.40)$$

est la fonction de répartition d'une loi normale  $\mathcal{N}(0, 1)$ .

#### Démonstration.

La différence d'énergie entre la labellisation correcte  $\lambda$  et une labellisation 1-optimale  $\lambda'$  est nulle partout sauf sur la région  $k_0$ , contenue dans le pixel  $y_0$ , donc

$$\begin{aligned} E(\lambda) - E(\lambda') &= \|V(y_0) - \sum_{k \in \mathcal{S}} \beta_k(y_0) m(\lambda_k)\|^2 - \|V(y_0) - \sum_{k \in \mathcal{S}} \beta_k(y_0) m(\lambda'_k)\|^2 \\ &= \|V(y_0) - \sum_{k \in \mathcal{S}} \beta_k(y_0) m(\lambda_k)\|^2 - \|V(y_0) - \sum_{k \in \mathcal{S}} \beta_k(y_0) m(\lambda_k) + \delta(y_0)\|^2, \end{aligned}$$

où  $\delta(y_0) = (\delta_t(y_0))_{t \in \mathcal{T}}$  et, pour tout  $t \in \mathcal{T}$ ,

$$\delta_t(y_0) = \beta_{k_0}(y_0) (m_t(\lambda_{k_0}) - m_t(\lambda'_{k_0})). \quad (5.4.41)$$

Pour toute date  $t \in \mathcal{T}$  et pour tout pixel  $y \in \mathcal{D}_{\mathbf{BR}}$ , définissons la variable aléatoire

$$\bar{V}_t(y) = V_t(y) - \sum_{k \in \mathcal{S}} \beta_k(y) m_t(\lambda_k). \quad (5.4.42)$$

Elle est gaussienne en tant que combinaison linéaire des variables aléatoires gaussiennes  $V_t(y)$ , de moyenne

$$\begin{aligned} \mathbb{E}[\bar{V}_t(y)] &= \mathbb{E}\left[V_t(y) - \sum_{k \in \mathcal{S}} \beta_k(y) m_t(\lambda_k)\right] \\ &= \mathbb{E}[V_t(y)] - \sum_{k \in \mathcal{S}} \beta_k(y) m_t(\lambda_k) \\ &= 0 \text{ d'après l'hypothèse (5.1.7),} \end{aligned} \quad (5.4.43)$$

et de variance

$$\begin{aligned}
 \mathbb{V} [\overline{V}_t(y)] &= \mathbb{V} \left[ V_t(y) - \sum_{k \in \mathcal{S}} \beta_k(y) m_t(\lambda_k) \right] \\
 &= \mathbb{V} [V_t(y)] \\
 &= \frac{\sigma^2}{N}.
 \end{aligned} \tag{5.4.44}$$

Le vecteur aléatoire  $\overline{V}(y) = (\overline{V}_t(y))_{t \in \mathcal{T}}$  est donc un vecteur gaussien centré et de matrice de covariance  $\frac{\sigma^2}{N} I_T$ , où  $I_T$  représente la matrice identité en dimension  $T$  (sous hypothèse d'indépendance des dates entre elles et d'égalité des variances des classes). Avec ces notations, la différence d'énergie entre les labellisations  $\lambda$  et  $\lambda'$  vaut alors

$$E(\lambda) - E(\lambda') = \|\overline{V}(y_0)\|^2 - \|\overline{V}(y_0) + \delta(y_0)\|^2. \tag{5.4.45}$$

La quantité  $\delta(y_0)$  étant constante, le vecteur aléatoire  $\overline{V}(y_0) + \delta(y_0)$  est gaussien, de moyenne  $\delta(y_0)$  et de matrice de covariance  $\frac{\sigma^2}{N} I_T$ . La probabilité qu'une labellisation 1-optimale fixée soit préférée à la labellisation correcte vaut alors

$$\begin{aligned}
 \mathbb{P}(E(\lambda) - E(\lambda') \geq 0) &= \mathbb{P}(\|\overline{V}(y_0)\|^2 - \|\overline{V}(y_0) + \delta(y_0)\|^2 \geq 0) \\
 &= \mathbb{P}(\|\overline{V}(y_0)\|^2 \geq \|\overline{V}(y_0) + \delta(y_0)\|^2) \\
 &= \mathbb{P}(\|\overline{V}(y_0)\|^2 \geq \|\overline{V}(y_0)\|^2 + 2 \langle \overline{V}(y_0), \delta(y_0) \rangle + \|\delta(y_0)\|^2) \\
 &= \mathbb{P}\left(\frac{2 \langle \overline{V}(y_0), \delta(y_0) \rangle}{\|\delta(y_0)\|^2} \leq -1\right).
 \end{aligned} \tag{5.4.46}$$

Or,

$$\mathbb{E}[\langle \overline{V}(y_0), \delta(y_0) \rangle] = \sum_{t \in \mathcal{T}} \delta^t(y_0) \mathbb{E}[\overline{V}^t(y_0)] = 0 \tag{5.4.47}$$

et

$$\begin{aligned}
 \mathbb{V}[\langle \overline{V}(y_0), \delta(y_0) \rangle] &= {}^t \delta(y_0) \mathbb{V}[\overline{V}(y_0)] \delta(y_0) \\
 &= {}^t \delta(y_0) \frac{\sigma^2}{N} I_T \delta(y_0) \\
 &= \frac{\sigma^2}{N} \|\delta(y_0)\|^2.
 \end{aligned} \tag{5.4.48}$$

Donc la variable aléatoire

$$\langle \overline{V}(y_0), \delta(y_0) \rangle \sim \mathcal{N}(0, \frac{\sigma^2}{N} \|\delta(y_0)\|^2) \tag{5.4.49}$$

et, en posant

$$X = \frac{2 \langle \overline{V}(y_0), \delta(y_0) \rangle}{\|\delta(y_0)\|^2}, \tag{5.4.50}$$

la variable aléatoire  $X$  suit une loi normale centrée, de variance  $\frac{4\sigma^2}{N\|\delta(y_0)\|^2}$ . Finalement, en notant  $Z = \frac{\sqrt{N}\|\delta(y_0)\|}{2\sigma}X$ , la variable aléatoire  $Z$  est centrée réduite, i.e.  $Z \sim \mathcal{N}(0, 1)$ .

La probabilité d'erreur sur la région  $k_0$ , d'après (5.4.46), s'écrit alors

$$\begin{aligned} \mathbb{P}(E(\lambda) - E(\lambda') \geq 0) &= \mathbb{P}(X \leq -1) \\ &= \mathbb{P}\left(Z \leq -\frac{\sqrt{N}\|\delta(y_0)\|}{2\sigma}\right). \end{aligned} \quad (5.4.51)$$

Soit  $G$  la fonction de répartition d'une loi normale  $\mathcal{N}(0, 1)$ , définie par

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt. \quad (5.4.52)$$

La probabilité qu'une labellisation 1-optimale soit préférée à la labellisation correcte vaut alors

$$\mathbb{P}(E(\lambda) - E(\lambda') \geq 0) = G\left(-\frac{\sqrt{N}\|\delta(y_0)\|}{2\sigma}\right), \quad (5.4.53)$$

où  $\|\delta(y_0)\| = \beta_{k_0}(y_0)\|m(\lambda_{k_0}) - m(\lambda'_{k_0})\|$ .  $\square$

Cette probabilité dépend du rapport de résolution  $N$  considéré, du taux d'occupation de la région erronée au sein du pixel basse résolution concerné, de l'erreur moyenne engendrée et de l'écart-type inter-classes  $\sigma$  (supposé égal pour chaque classe). Le corollaire 5.4.3 donne un encadrement de la probabilité en supposant que toutes les régions de la segmentation ne sont pas trop petites relativement au pixel basse résolution.

**Corollaire 5.4.3** *Si toutes les régions de la segmentation occupent une surface en proportion supérieure à  $n/N$  d'un pixel basse résolution, alors la probabilité de préférer une labellisation 1-optimale  $\lambda'$  à la labellisation correcte  $\lambda$  vérifie les deux inégalités suivantes :*

$$\mathbb{P}((E(\lambda) - E(\lambda')) \geq 0) \leq G\left(-\frac{\sqrt{N}}{2} \frac{n}{N} \frac{\min_{\lambda \neq \lambda'} \|m(\lambda) - m(\lambda')\|}{\sigma}\right), \quad (5.4.54)$$

et

$$\mathbb{P}((E(\lambda) - E(\lambda')) \geq 0) \geq G\left(-\frac{\sqrt{N}}{2} \frac{\max_{\lambda \neq \lambda'} \|m(\lambda) - m(\lambda')\|}{\sigma}\right). \quad (5.4.55)$$

**Démonstration.**

D'après la proposition 5.4.2,

$$\mathbb{P}((E(\lambda) - E(\lambda')) \geq 0) = G\left(-\frac{\sqrt{N}\beta_{k_0}(y_0)\|m(\lambda_{k_0}) - m(\lambda'_{k_0})\|}{2\sigma}\right). \quad (5.4.56)$$

Si toutes les régions de la segmentation impactent un pixel basse résolution en proportion supérieure à  $\frac{n}{N}$ , alors

$$\beta_{k_0}(y_0)\|m(\lambda_{k_0}) - m(\lambda'_{k_0})\| \geq \frac{n}{N} \min_{\lambda \neq \lambda'} \|m(\lambda_{k_0}) - m(\lambda'_{k_0})\|. \quad (5.4.57)$$



De plus, la fonction  $x \mapsto G(x)$  est une fonction croissante de  $x$ , donc

$$\mathbb{P}((E(\lambda) - E(\lambda')) \geq 0) \leq G\left(-\frac{\sqrt{N}}{2} \frac{n}{N} \frac{\min_{\lambda \neq \lambda'} \|m(\lambda) - m(\lambda')\|}{\sigma}\right). \quad (5.4.58)$$

Inversement, quelque soient  $k_0$  et  $y$  fixés,  $\beta_{k_0}(y) \leq 1$  donc

$$\beta_{k_0}(y_0) \|m(\lambda_{k_0}) - m(\lambda'_{k_0})\| \leq \max_{\lambda \neq \lambda'} \|m(\lambda) - m(\lambda')\|,$$

donc

$$\mathbb{P}((E(\lambda) - E(\lambda')) \geq 0) \geq G\left(-\frac{\sqrt{N}}{2} \frac{\max_{\lambda \neq \lambda'} \|m(\lambda) - m(\lambda')\|}{\sigma}\right). \quad (5.4.59)$$

□

Ce résultat donne un intervalle de fiabilité en fonction de la *taille* des pixels basse résolution, de la taille minimale d'une région relativement à un pixel basse résolution et du rapport  $\frac{\min_{\lambda \neq \lambda'} \|m(\lambda) - m(\lambda')\|}{\sigma}$ . Remarquons que ce dernier rapport s'apparente à une mesure de *contraste*, on note

$$c = \frac{\min_{\lambda \neq \lambda'} \|m(\lambda) - m(\lambda')\|}{\sigma}. \quad (5.4.60)$$

La figure 5.1 trace les valeurs de ce majorant en fonction du *contraste* de l'image, pour des régions de taille minimale  $\frac{n}{N}$  (avec différentes valeurs de  $\frac{n}{N}$ ). Remarquons notamment qu'une labellisation 1-optimale réalisée à partir d'une image de *contraste* supérieur à 1 a une probabilité infime d'être préférée à la labellisation correcte si toutes les régions de la segmentation représentent une proportion supérieure à 30 du pixel **BR**. De plus, la probabilité de faire une erreur de labellisation sur une région contenue dans un pixel basse résolution est d'autant plus faible que le *contraste* est fort.

Le corollaire 5.4.3 soulève le problème de la taille des régions d'une segmentation relativement à l'échelle d'observation. Dans la section 6.6, nous présentons une analyse empirique de la sensibilité de la méthode au facteur de résolution. A ce titre, nous évoquons la question de la taille minimale des régions distinguables.

Par ailleurs, dans l'hypothèse où l'utilisateur souhaite se restreindre à l'utilisation de  $\theta$  dates de l'ensemble  $\mathcal{T}$ , le résultat 5.4.2 dicte une manière de les choisir. En effet, bien que le calcul de la probabilité d'erreur ne soit fait que dans le cas précis où seule une région est erronée, il donne une idée des paramètres qui influent sur les performances. En particulier, remarquons que le *contraste*, défini par la distance minimale entre les moyennes des classes normalisée par la variance, a un effet non négligeable sur la probabilité de mal labelliser une région. Une manière de choisir  $\theta$  dates dans  $\mathcal{T}$  tout en minimisant les erreurs de labellisation serait alors de résoudre le problème

$$\max_{t_1 < \dots < t_\theta} \min_{\lambda \neq \lambda'} \sum_{i=1}^{\theta} |m_{t_i}(\lambda) - m_{t_i}(\lambda')|^2. \quad (5.4.61)$$

La recherche d'une solution au problème (5.4.61) peut se faire en explorant tous les sous-ensembles de  $\theta$  parmi  $T$ , mais cette exploration de  $C_T^\theta$  possibilités est rapidement très coûteuse en temps

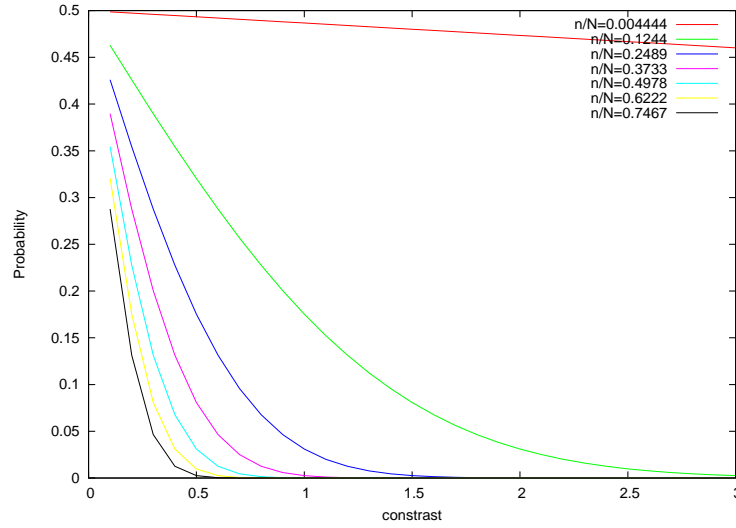


FIG. 5.1.: Majorant de la probabilité de préférer une labellisation 1-optimale à la labellisation correcte lorsque la région mal labellisée représente une proportion supérieure à  $n/N$  d'un pixel basse résolution. La valeur du majorant est tracée en fonction du niveau *contraste* de l'image. L'axe des abscisses représente le niveau de contraste  $c$  et l'axe des ordonnées le majorant de la probabilité. Chaque courbe correspond à une proportion  $n/N$  fixée, allant de 0.4% à 74.6%.

de calcul. Dans cette thèse, nous ne développons pas ce point mais il pourra faire l'objet de travaux complémentaires dans la mesure où le choix des dates d'acquisition constitue un problème important pour la qualité d'une classification ainsi que pour le temps de calcul.

Dans le chapitre 6, nous décrivons un algorithme de recuit simulé pour résolution du problème (5.2.19). Il permet d'obtenir une labellisation et d'estimer les moyennes des classes à partir d'une séquence d'images **BR** et d'une segmentation **HR**. La convergence de cet algorithme est ensuite étudiée de manière expérimentale avant d'analyser la sensibilité de la méthode en fonction du facteur de résolution.



## 6. Algorithmes et performances empiriques

Ce chapitre est centré sur la mise en oeuvre du modèle bayésien décrit dans le chapitre 5 pour la classification HR de l'occupation du sol à partir d'une séquence d'images BR et d'une segmentation HR. Dans un premier temps, nous décrivons l'algorithme de recuit simulé utilisé pour trouver une solution. Puis nous analysons les performances de l'algorithme et de la méthode vis-à-vis des principaux paramètres considérés.

### 6.1. Algorithme de recuit simulé

La minimisation des énergies (5.2.19) ou (5.2.16) est un problème combinatoire difficile à résoudre. En effet, compte tenu de la taille de l'espace des solutions ( $L^S$ ), une recherche systématique du minimum n'est pas envisageable. Nous proposons d'adopter un algorithme de type recuit simulé pour résoudre le problème 5.2.19, *i.e.*

$$\lambda^* = \arg \min_{\lambda \in L^S} \sum_{t \in T} \sum_{y \in \mathcal{D}_{BR}} (v_t(y) - \mu_t(\lambda, y))^2,$$

où

$$\mu_t(\lambda, y) = \sum_{l \in \mathcal{L}} \sum_{\substack{k \in \mathcal{S} \\ l_k = l}} \beta_k(y) m_t(l).$$

Les algorithmes de recuit simulé sont inspirés du recuit en physique statistique qui consiste à chauffer (jusqu'à fusion) puis laisser refroidir un métal dans le but d'améliorer ses qualités, et ce suffisamment lentement pour ne pas le *bloquer* dans un état instable correspondant à un minimum local d'énergie. L'utilisation du recuit simulé est destinée aux problèmes de minimisation difficiles (*cf.* [Ycart, 2002]), *i.e.* lorsqu'une recherche systématique est rendue impossible par la taille de l'espace, et lorsque la fonction à minimiser a un grand nombre de minima locaux qui perturbent l'accès au minimum global. [Geman et Geman, 1984] ont été les premiers à utiliser un algorithme de recuit simulé en traitement d'images, pour minimiser ce type d'énergies. Le principe de l'algorithme est proche de celui d'une descente de gradient, dont les pas de descente sont calculés automatiquement en fonction de la température et pour lequel des remontées d'énergie sont autorisées avec une certaine probabilité.

Dans un contexte non-supervisé, l'algorithme prend en entrée une segmentation **HR**, une séquence d'images **BR** de la même scène, et le nombre de labels  $L$  recherché. Il permet d'obtenir, en sortie, une labellisation qui correspond à une solution approchée du problème (5.2.19) ainsi

que l'estimation des moyennes caractéristiques des classes associées à la labellisation obtenue. A partir d'une initialisation aléatoire de la labellisation et de la température initiale calculée à partir du diamètre du graphe, chaque étape de l'algorithme consiste à modifier aléatoirement le label d'une région de la segmentation puis à calculer la différence d'énergie engendrée de manière à déterminer si cette modification fait décroître l'énergie ou non. Dans un contexte non-supervisé, nous utilisons l'algorithme 1. Pour considérer le problème (5.2.16), *i.e.* sans hypothèse d'égalité des variances des classes, il suffit d'insérer l'estimation des variances des classes après chaque estimation des moyennes des classes dans l'algorithme 1. En pratique, nous préférons nous restreindre aux statistiques du premier ordre en raison du coût en temps de calcul important engendré par ces estimations dans un algorithme itératif.

- Calculer,  $\forall y \in \mathcal{D}_{BR}$ , les proportions des régions  $(\beta_k(y))_{y,k}$ ,
- Initialiser aléatoirement le champ des labels,  $\lambda$ ,
- Estimer les moyennes associées à chaque label (par régression linéaire),
- Calculer l'énergie  $\mathcal{E}_\lambda$  associée à la labellisation  $\lambda$ ,
- Initialiser la température  $\theta$  au diamètre du graphe,
- Initialiser le nombre de tests rejetés successivement,  $n_r = 0$ ,
- Tant que  $n_r \leq a \times S$  fois, faire :
  - pour  $i = 0$  à  $S$ ,
  - sélectionner aléatoirement une région  $k$  et un label  $l$  pour cette région,
  - ré-estimer les moyennes associées à chaque label,
  - calculer la différence d'énergie  $\Delta\mathcal{E} = \mathcal{E}_\lambda - \mathcal{E}_{\lambda_{prev}}$ , où  $\lambda_{prev}$  est la labellisation obtenue à l'itération précédente,
  - si  $\Delta\mathcal{E} \leq 0$ , accepter le changement de label,  $n_r = 0$
  - sinon, le rejeter avec une probabilité  $\exp(-\Delta\mathcal{E}/\theta)$ ,  $n_r++$ .
- $\theta \leftarrow q\theta$ .

**Algorithme 1:** Algorithme de recuit simulé pour la labellisation d'une segmentation **HR** à partir d'une séquence d'images **BR**.

D'un point de vue théorique, la convergence des algorithmes de recuit est très étudiée dans la littérature. En particulier, le théorème de Hajek (1988) prouve, sous certaines conditions, la convergence d'un algorithme de recuit simulé pour un schéma de température correspondant à une décroissance logarithmique. De plus, des résultats théoriques tels que [Catoni, 1998] discutent du problème du test d'arrêt. En pratique, compte tenu de la taille des images considérées et du nombre de régions de la segmentation, une descente logarithmique n'est pas envisageable pour des questions de temps de calcul. En revanche, une décroissance géométrique de raison  $q = 0.999$  et un test d'arrêt fixé par  $n_r = 400$  ont montré, empiriquement, de bons résultats dans les expériences réalisées. Dans la section 6.5, une série d'expériences nous permet d'analyser les erreurs de labellisation obtenues dans le but de déterminer si elles proviennent d'un problème de convergence de l'algorithme ou de la modélisation.

Lorsque les moyennes caractéristiques des classes sont connues *a priori*, une version supervisée de l'algorithme s'obtient directement à partir de l'algorithme 1, non-supervisé, en supprimant les phases d'estimation et de ré-estimation des moyennes associées à chaque label. La version supervisée prend alors en entrée une segmentation haute résolution, une séquence d'images **BR** et les caractéristiques des classes. Dans ce cas, l'algorithme peut directement résoudre le problème (5.2.16) en considérant l'énergie qui lui est associée. En sortie, il fournit la labellisation optimale au sens du maximum *a posteriori*. L'estimation des moyennes des classes, à chaque itération, est une tâche coûteuse en temps de calcul. La suppression de cette étape conduit donc à un algorithme nettement plus rapide que la version non-supervisée.

## 6.2. Algorithme de programmation dynamique

La question de la validation d'une méthode, pour ce type d'applications, est une question délicate puisqu'une carte de référence n'est pas toujours accessible et, même si elle l'est, elle n'est pas toujours fiable. L'objectif que nous visons est, par conséquent, d'être capable d'obtenir une labellisation à partir d'images **BR** qui soit aussi proche que possible de ce qu'on obtiendrait avec des images **HR**.

L'algorithme de recuit simulé présenté dans la section 6.1 peut être appliqué directement à des images ou séquences d'images **HR**. La solution à laquelle il permet d'accéder est une estimation du minimum de l'énergie dans un contexte stochastique. Pour l'élaboration d'une classification de référence, nous proposons dans cette section un algorithme de programmation dynamique capable de déterminer précisément une solution à partir d'images **HR**.

A haute résolution, les pixels d'une image sont considérés comme *purs* : chaque pixel ne représente qu'un seul type d'occupation. A partir d'une segmentation et d'une séquence de même résolution, l'algorithme de programmation dynamique 2 permet alors de regrouper les  $S$  régions autour de  $L$  labels, pour un nombre de labels  $L$  fixé (cf. figure 6.1). Cet algorithme est déterministe et garantit d'accéder au minimum de l'énergie.

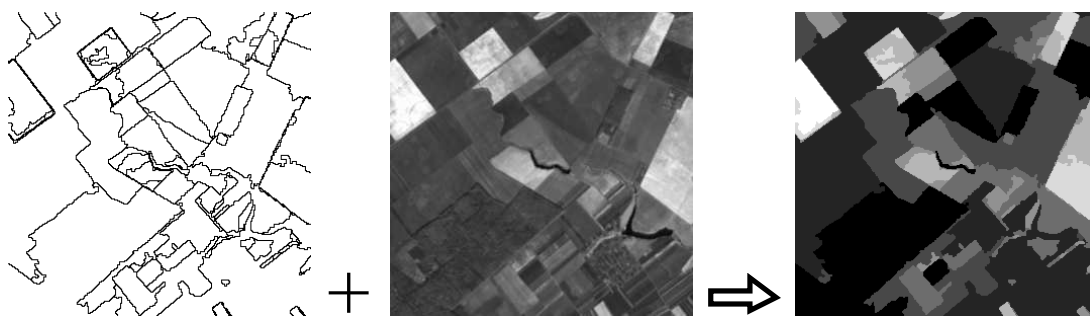


FIG. 6.1.: Entrées et sortie de l'algorithme : à partir d'une segmentation (300 régions) et d'une image de même résolution, l'algorithme permet d'obtenir une classification pour un nombre de labels fixé (ici,  $L = 8$ ).

A partir d'une segmentation contenant  $S$  régions et d'une image observée, la valeur radiométrique moyenne de chaque région s'obtient directement. Pour toute région, notons  $m_i$  la moyenne associée à la région  $i$  et  $N_i$  sa taille (en nombre de pixels). Le principe de l'algorithme repose sur le tri des régions selon leur valeur moyenne. Supposons, par conséquent, que les régions soient indicées selon leur valeur moyenne, par ordre croissant :  $m_1 \leq m_2 \leq \dots \leq m_n$ . On note  $\mathcal{S}_l$  l'ensemble des moyennes de toutes les régions de label  $l$ . L'objectif est alors de regrouper ces  $S$  régions autour de  $L$  labels ( $L < S$ ).

A haute résolution, l'énergie considéré par le problème (5.2.19) peut s'écrire sous la forme

$$\mathcal{E}(\lambda) = \sum_{l \in \mathcal{L}} \sum_{k; l_k=l} \sum_{x; s(x)=k} (u(x) - \mu_l)^2, \quad (6.2.1)$$

où  $\mu_l = \frac{\sum_k N_k m_k}{\sum_k N_k}$ . En développant, cette énergie s'écrit aussi sous la forme

$$\begin{aligned} \mathcal{E}(\lambda) &= \sum_x u(x)^2 + \sum_l \sum_k (-2N_k m_k \mu_l + N_k \mu_l^2) \\ &= \sum_x u(x)^2 + \sum_l \sum_k N_k (m_k - \mu_l)^2 - \sum_k N_k m_k^2, \end{aligned} \quad (6.2.2)$$

où  $\sum_x u(x)^2$  et  $\sum_k N_k m_k^2$  sont des termes constants. Pour une famille de moyennes  $\mathcal{S}_l$ , associées au label  $l$ , on note

$$\Sigma_\lambda(\mathcal{S}_l) = \sum_k N_k (m_k - \mu_l)^2. \quad (6.2.3)$$

L'énergie (6.2.1) s'écrit alors

$$\sum_{l \in \mathcal{L}} \Sigma_\lambda(\mathcal{S}_l). \quad (6.2.4)$$

Avec ces notations, si les régions sont indicées par moyennes croissantes, la labellisation qui minimise (6.2.1) s'obtient directement en regroupant des régions d'indices consécutifs. Plus précisément, l'objectif est de trouver l'application  $f : \{1, \dots, L\} \rightarrow \{1, \dots, S\}$  croissante, vérifiant

$$f(1) = 1, f(L) = S, \quad (6.2.5)$$

et telle que  $\sum_{i=1}^L \Sigma_\lambda(\mathcal{S}_i)$  soit minimale. Cette fonction permet de décrire l'ensemble  $\mathcal{S}_i$  des moyennes des régions de label  $i$  par

$$\mathcal{S}_i = \{m_{f(i)}, m_{f(i)+1}, \dots, m_{f(i+1)-1}\}. \quad (6.2.6)$$

Il s'agit alors de déterminer les indices  $f(l)$ , où  $l = \{1, \dots, L\}$ , qui indiquent la première région du label  $l$  dans le vecteur ordonné. Pour cela, considérons les fonctions définies, pour  $k \in \{1, \dots, S\}$  et  $l \in \{1, \dots, L\}$ , par

$$\varphi(k, l) = \min_{k=f(l) \leq \dots \leq f(L)=n} \sum_{i=l}^L \Sigma_\lambda(\mathcal{S}_i), \quad (6.2.7)$$

et

$$\psi(k, l) = \arg \min_{k=f(l) \leq \dots \leq f(L)=n} \sum_{i=l}^L \Sigma_{\lambda}(\mathcal{S}i). \quad (6.2.8)$$

La programmation dynamique repose sur le fait que la fonction  $\varphi$  s'écrit de manière récursive (équation de Bellman). En effet, pour tout  $k \in \{1, \dots, n\}$ ,

$$\varphi(k, l) = \min_{i \in [k+1, n]} \{ \Sigma_{\lambda}\{m_k, m_{k+1}, \dots, m_{i-1}\} + \varphi(i, l+1) \}. \quad (6.2.9)$$

Dans le cas où  $l = L$ , pour tout  $k \in \{1, \dots, n\}$ , on obtient

$$\psi(k, L) = \arg \min_{f; k=f(l)} \sum_{i=L}^L \Sigma_{\lambda}(\mathcal{S}i) \quad (6.2.10)$$

$$= \Sigma_{\lambda}\{m_k, m_{k+1}, \dots, m_n\}. \quad (6.2.11)$$

Pour tout  $k$ , ces relations permettent de déterminer  $\varphi(k, L-1)$  à partir de  $\varphi(k, L)$  et, récursivement, tous les termes jusqu'à  $\varphi(k, 1)$ . L'application  $f$  recherchée correspond alors à l'argument du minimum de  $\varphi(1, 1)$ , i.e.  $\psi(1, 1)$ . Etant donnée une segmentation contenant  $S$  régions et une image **HR**  $u$ , l'algorithme 2 décrit la recherche de la labellisation  $\lambda$  qui correspond au minimum de l'énergie.

- $\forall k \in \mathcal{S}$ , calculer la taille  $N_k$  de la région  $k$ .
- $\forall k \in \mathcal{S}$ , calculer la moyenne  $m_k$  de la région  $k$ .
- Trier le vecteur  $(m_k)_{k \in \mathcal{S}}$  selon l'ordre croissant,  $p_k$  désigne alors la région de rang  $k$ .
- $\forall (i, j) \in \mathcal{S}^2$ , calculer  $var(i, j) = \sum_{k=i}^j N_k \left( m_k - \frac{\sum_{l=i}^j N_l m_l}{\sum_{l=i}^j N_l} \right)^2$ .
- $\forall k \in \mathcal{S}$ , calculer  $\varphi(k, L) = var(k, S)$ .
- Pour  $l = (L-1) \dots 1$ 
  - Pour  $k = (S-1) \dots 1$ 
    - calculer  $\varphi(k, l) = \min_{i \in \{k+1, \dots, S-1\}} var(k, i-1) + \varphi(i, L+1)$ ,
    - $\psi(k, l) = \arg \min_{i \in \{k+1, \dots, S-1\}} var(k, i-1) + \varphi(i, L+1)$ .
- Initialiser  $f(0) = 0$ .
- Pour  $l = 1 \dots (L-1)$ ,  $f(l) = \psi(f(l), l)$  et  $f(l) = S$ .
- Pour  $l = 1 \dots L$ , pour  $j = f(l) \dots f(l+1) - 1$ , le label de la région  $p_j$  vaut  $l$ .

**Algorithme 2:** Algorithme de programmation dynamique pour la labellisation optimale d'une segmentation **HR** à partir d'une image **HR**.

Cet algorithme permet de déterminer le minimum de l'énergie (6.2.1) de manière exacte, à partir d'une image haute résolution. Pour les expériences qui suivent, cette labellisation est considérée en tant que référence et les résultats obtenus à partir d'images **BR** sont comparés à ceux qu'on obtiendrait avec des images **HR**.



### 6.3. Simulation des données

Dans cette section, nous décrivons la manière dont les images sont simulées pour évaluer la méthode de classification proposée et, en particulier, permettre une analyse quantitative et qualitative des différents types d'erreurs et des performances.

Pour respecter la structure globale des images en observation de la Terre ainsi que leur dynamique, les images ont été simulées à partir de caractéristiques extraites sur des images réelles. Plus précisément, nous avons constitué un jeu de données de 165 images de taille  $300 \times 300$  pixels extraits de la séquence d'images réelles **HR** de la série ADAM<sup>1</sup> (Plaine du Danube, Roumanie). Ces images ont été segmentées arbitrairement en 100 régions ( $S = 100$ ), avec le module `segct` du logiciel <http://cmla.ens-cachan.fr/Cmla/Megawave>, de manière à respecter globalement la structure spatiale de l'image, puis labellisées pour  $L = 5$  labels avec l'algorithme de programmation dynamique (cf. 6.2) en vue de constituer des labellisations de référence.

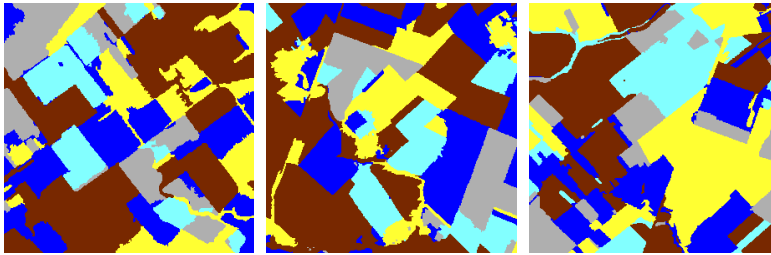


FIG. 6.2.: Exemples de labellisations simulées avec 5 labels.

Les moyennes et variances associées à chaque label sont alors utilisées pour simuler des données par mélange de gaussiennes : pour chaque labellisation de référence, une image **HR** est simulée en tirant chaque pixel suivant une loi gaussienne paramétrée selon les moyennes et variances réelles associées au label correspondant. Les images **BR** sont alors créées par moyennage des images **HR** simulées (avec les facteurs de résolution  $N = 15, 30, 50$ .)

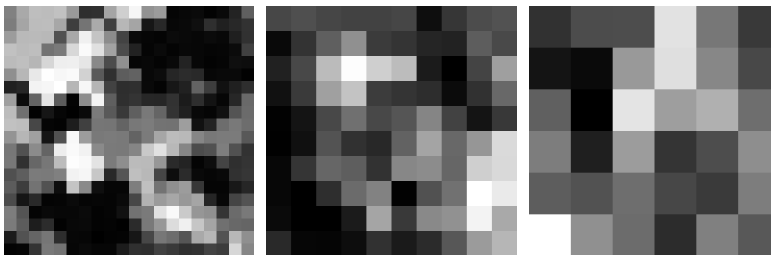


FIG. 6.3.: Images **BR** simulées à partir des labellisations ci-dessus, avec facteurs de résolution, de gauche à droite,  $N = 15, 30, 50$ .

---

<sup>1</sup>Projet européen d'agro-modélisation. Images mises à disposition par le CNES.

Avant d'analyser empiriquement les performances de la méthode sur un grand jeu de données simulées, la section suivante présente quelques résultats de labellisation obtenus à partir de séquences d'images **BR** simulées. En particulier, elle illustre la prise en compte de l'information spatiale pour la labellisation.

## 6.4. Exemples d'application

Dans cette section, nous présentons les premiers résultats obtenus sur des séquences d'images simulées (*cf.* section 6.3) dans l'objectif d'illustrer la capacité de la méthode de classification à exploiter l'information temporelle.

Les premières expériences ont été réalisées dans l'objectif de présenter l'apport d'informations redondantes pour la qualité de la labellisation. Pour cela, trois images sont simulées selon les mêmes caractéristiques (moyennes et variances par label) mais les labels auxquels ces caractéristiques ont été associées ont été permutés entre les différentes images de manière ce que la séquence d'images contienne, pour chaque label, des informations redondantes. La figure 6.4 montre, sur un exemple donné, que les erreurs de labellisation obtenues sont de moins en moins nombreuses lorsqu'on augmente le nombre d'images utilisées. Cette expérience permet d'illustrer l'apport d'informations redondantes sur la qualité de la labellisation.

La figure 6.5 présente la même expérience réalisée avec des images simulées de manière à ce que chaque image comporte deux classes confondues. Le pourcentage de pixels mal labellisés passe de 1.12% lorsque deux images sont utilisées à 0.09% pour trois images. La complémentarité des informations apportées par une séquence d'images favorise nettement la distinction des classes.

L'expérience suivante illustre l'intérêt d'informations complémentaires pour la labellisation. Les images sont simulées de telle sorte que plusieurs classes soient confondues à une date donnée mais que l'ensemble des classes soit séparable sur l'ensemble des dates. Les profils d'évolutions temporelles des moyennes utilisés sont présentés figure 6.6 en (a), et les images (b à f) correspondent aux images **BR** obtenues par moyennage d'un facteur 15 des images simulées par tirage gaussien respectivement selon les moyennes caractéristiques associées aux dates 1 à 5. La labellisation obtenue avec ces 5 images est présentée figure 6.6 (i). Elles contiennent 5 régions mal labellisées, qui correspondent à seulement 0.28% des pixels de l'image.

Les expériences présentées dans cette section confirment, d'un point de vue qualitatif, l'importance de l'information temporelle pour discriminer les différents types d'occupation. Si l'apport d'informations redondantes contribue à distinguer les classes, l'information complémentaire apportée à chaque date est essentielle dans le contexte du suivi des surfaces continentales. En effet, l'intensité associée à plusieurs types de végétation étant généralement confondue à une date fixée, seule son évolution temporelle permet de les séparer. Dans la section suivante, nous nous intéressons aux différents types d'erreurs de labellisation que l'on peut obtenir avec la méthode non-supervisée : les erreurs dues à la non-convergence de l'algorithme et celles propres à l'énergie considérée.

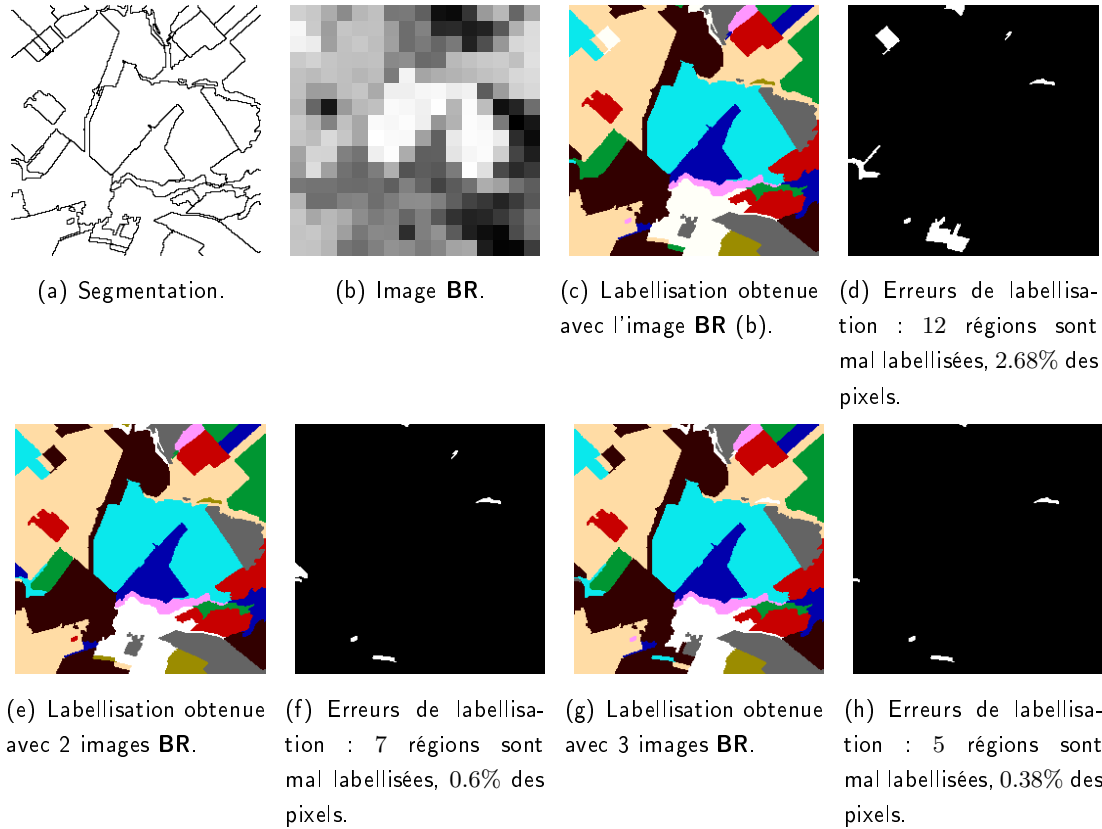


FIG. 6.4.: Labellisation pour 10 labels de la segmentation présentée en (a) à partir de l'image **BR** (b) (rapport de résolution  $N = 15$ ), puis à partir d'une séquence basse résolution contenant 2 images ((b) et une autre image similaire), et enfin à partir de 3 images. Les erreurs de labellisation diminuent lorsqu'on augmente le nombre d'images exploitées.

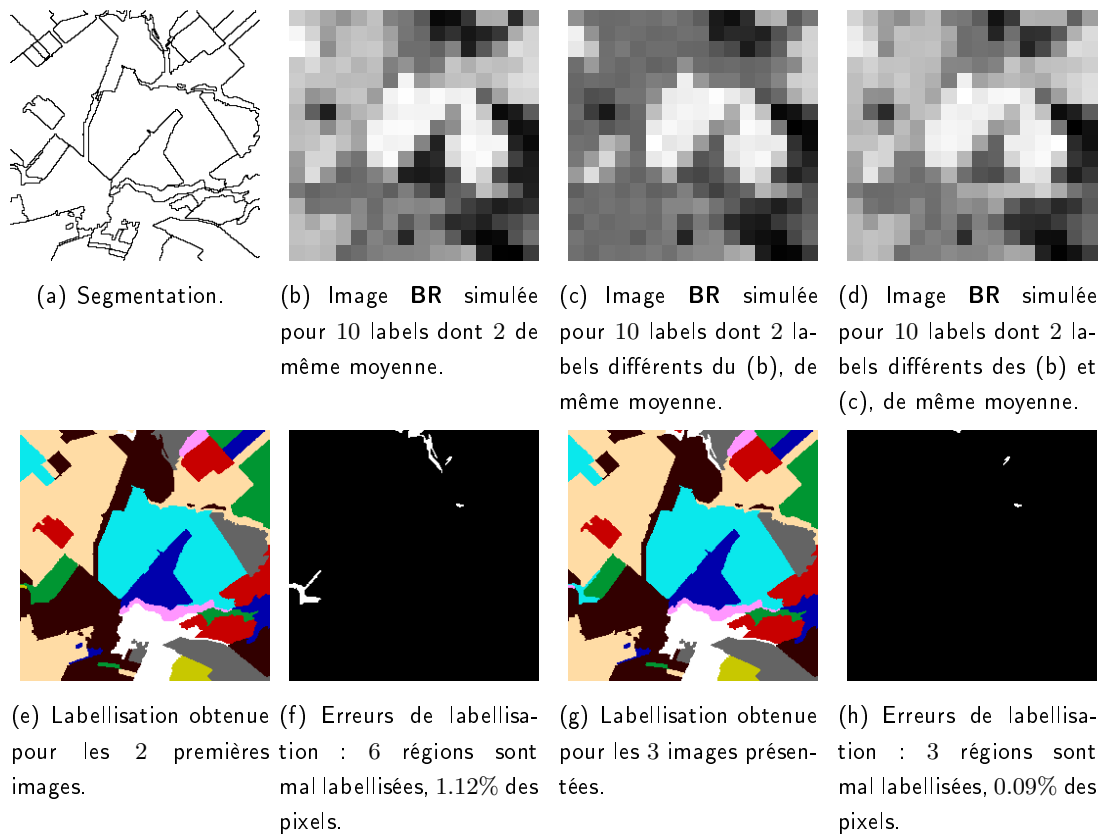
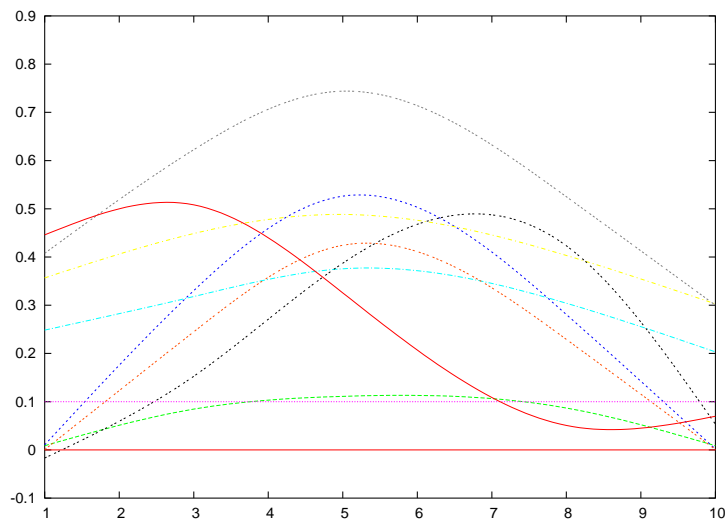
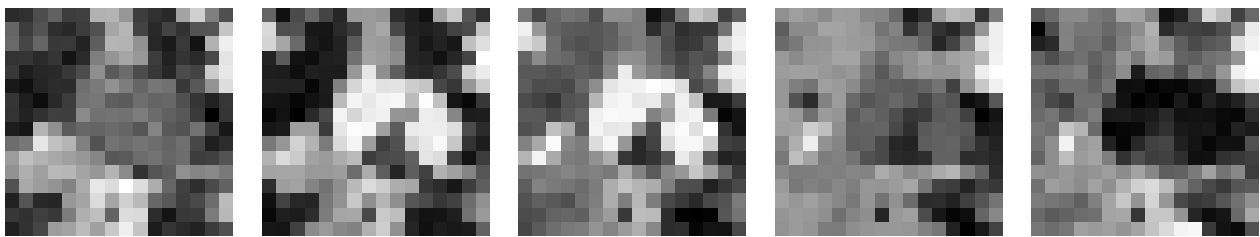


FIG. 6.5.: Labellisation pour 10 labels de la segmentation présentée en (a) à partir de la séquence **BR** composée des images (b) et (c) (rapport de résolution  $N = 15$ ), puis à partir des images (b), (c) et (d). Les erreurs de labellisation diminuent avec lorsque des informations redondantes sont utilisées.



(a) Evolutions temporelles des moyennes utilisées pour simuler la séquence d'images **BR** : chaque courbe correspond à un label différent. L'échelle de temps est représenté en abscisses et les valeurs moyennes des classes (fraction de couvert) en ordonnées.



(b) Date 1.

(c) Date 2.

(d) Date 3.

(e) Date 4.

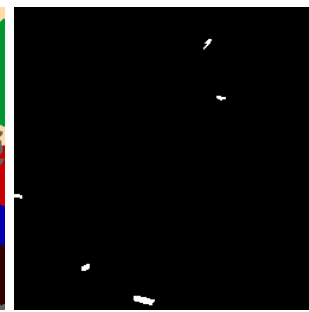
(f) Date 5.



(g) Segmentation **HR**.



(h) Labellisation obtenue pour la séquence **BR** composée d'images (b) à (f).



(i) En blanc, erreurs de labellisation : 5 régions sont mal labellisées, 0.28% des pixels.

FIG. 6.6.: Labellisation obtenue en mode non-supervisé, avec une séquence de 5 images **BR**, simulées à partir des profils temporels (a) et avec un facteur de résolution de 15 par rapport à la résolution de la segmentation (g). La carte des erreurs (i) montre que 5 régions ont été mal labellisées, elles représentent 0.28% des pixels de l'image.

## 6.5. Analyse des erreurs

Dans cette section, nous limitons l'étude au cas monotemporel puisque notre objectif est simplement de tester la convergence de l'algorithme de recuit simulé en analysant les différents types d'erreurs obtenus. Le tableau 6.1 présente les résultats obtenus globalement par les approches supervisées et non-supervisées sur un jeu de 165 images simulées.

TAB. 6.1.: Résultats obtenus pour l'ensemble des 165 images simulées avec un rapport de résolution  $\frac{|\mathcal{D}_{\text{HR}}|}{|\mathcal{D}_{\text{BR}}|} = 15 \times 15$ .

	Pixels mal labellisés (% moyen)	Régions mal labellisées (% moyen)
Version supervisée	0.87	23.6
Version non-supervisée	4.35	31.5

Les résultats obtenus pour la méthode supervisée, sur les 165 images simulées, sont très satisfaisants puisque moins de 1% des pixels, en moyenne, ont été mal labellisés. Si le pourcentage moyen de régions mal labellisées, quant à lui, semble très élevé, il résulte simplement de la présence de petites régions dans la segmentation puisque le pourcentage de pixels concernés est très faible. Pour éviter ce type d'artéfacts, un seuil sur la taille des régions pourrait être imposé lors de la segmentation. Ce type de pré-traitement était déjà suggéré par le corollaire 5.4.3 qui donne une majoration de la probabilité d'erreur de labellisation en fonction de ce seuil et du niveau de contraste de l'image. Bien que les erreurs de labellisation soient, en moyenne, plus nombreuses avec la version non-supervisée, elles restent raisonnables avec moins de 5% des pixels mal labellisés. Les résultats restent donc plutôt convaincants, sachant que la valeur médiane des erreurs est de 2%.

Le premier objectif de cette analyse est de mettre en évidence les erreurs qui résultent d'un problème d'optimisation, montrant que l'algorithme n'a pu atteindre le minimum global. Pour limiter les sources d'erreurs, nous nous plaçons dans un contexte supervisé et nous considérons l'énergie définie par la proposition 5.2.2 :

$$\mathcal{E}_\lambda(v) = \sum_{t \in \mathcal{T}} \sum_{y \in \mathcal{D}_{\text{BR}}} \left( \frac{(v_t(y) - \mu_t(\lambda, y))^2}{\sigma_t^2(\lambda, y)} + \ln(\sigma_t^2(\lambda, y)) \right). \quad (6.5.12)$$

Pour séparer les erreurs d'optimisation des autres types d'erreurs, nous comparons la valeur de l'énergie, notée  $\mathcal{E}_0(v)$ , associée à la labellisation de référence (pour une image  $v$  fixée) à la valeur de l'énergie, notée  $\mathcal{E}_k(v)$ , associée à une labellisation qui serait égale en tout point à la labellisation de référence sauf sur la région  $k$ . Nous calculons la différence  $\mathcal{E}_k(v) - \mathcal{E}_0(v)$  successivement pour toutes les régions  $k$  mal labellisées par l'algorithme.

A titre de comparaison, nous nous intéressons à la distribution des variations d'énergie qu'engendrerait l'attribution d'un mauvais label à une région. La figure 6.7 (a) présente l'histo-

gramme des variations d'énergies normalisées  $((\mathcal{E}_k(v) - \mathcal{E}_0(v))/\mathcal{E}_0(v))$  calculées pour toutes les régions de  $\mathcal{S}$  labellisées successivement avec tous les labels, à l'exception du label correct. Pour son interprétation, il est nécessaire de préciser que, en pratique,  $\mathcal{E}$  admet des valeurs négatives. Si la plus petite valeur de différence d'énergie obtenue est  $-1567$ , pour plus de lisibilité la figure 6.7 ne présente que les occurrences majoritaires, situées entre  $-8$  et  $0.1$  (en fait, elles sont surtout concentrées entre  $-2$  et  $0.1$ ). Remarquons que presque 50% des occurrences se réfèrent à des différences d'énergie comprises entre  $-0.1$  et  $0$ . De plus, remarquons la présence de valeurs positives, inférieures à  $0.004$ , pour environ 10% des occurrences, cas que nous étudierons plus loin.

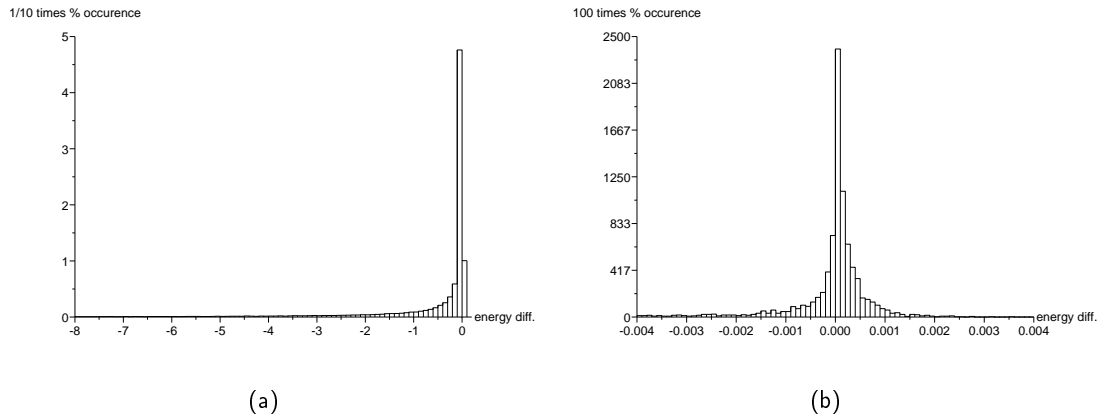


FIG. 6.7.: (a) Histogramme de la différence d'énergies normalisée  $(\mathcal{E}_k(v) - \mathcal{E}_0(v))/\mathcal{E}_0(v)$  calculée pour toutes les régions de la segmentation labellisées successivement avec tous les labels incorrects. (b) Histogramme de la différence d'énergies normalisée  $(\mathcal{E}_k(v) - \mathcal{E}_0(v))/\mathcal{E}_0(v)$  calculée pour toutes les erreurs obtenues avec l'algorithme supervisé. La comparaison de ces histogrammes montre, notamment, que les erreurs de labellisation dues à un problème de convergence sont associées à des différences d'énergies très faibles par rapport à celles qu'on obtiendrait avec d'autres labels inexacts).

La figure 6.7 (b) représente l'histogramme des différences d'énergies calculées pour toutes les labellisations contenant exactement une région mal labellisée parmi toutes les labellisations incorrectes obtenues. Les valeurs négatives correspondent typiquement à un problème de convergence de l'algorithme. En effet,  $\mathcal{E}_0$  étant négative,  $(\mathcal{E}_k(v) - \mathcal{E}_0(v))/\mathcal{E}_0(v)$  est négative si  $\mathcal{E}_k(v) \geq \mathcal{E}_0(v)$ . L'attribution du label incorrect considéré a donc bien engendré une croissance de l'énergie mais l'algorithme n'a pas été capable de sortir du minimum local pour atteindre le minimum global. Ce problème de convergence de l'algorithme représente 41.3% des occurrences. Cependant, remarquons que les valeurs des différences d'énergies obtenues sont nettement plus faibles que celles obtenues dans le cas de la figure 6.7 (a) (environ 1000 fois plus faibles). Les erreurs de convergence sont, par conséquent, limitées au cas de minima locaux très proches des minima globaux. Une amélioration de l'algorithme pourrait être obtenue en modifiant le pas ou la loi de descente en température (pour un pas plus fin ou une loi logarithmique). Toutefois, il est nécessaire de faire un compromis

entre temps de calcul et performance des résultats.

De plus, les différences d'énergie positives qui apparaissent dans l'histogramme (figure 6.7 (b)) correspondent à des occurrences où la valeur  $\mathcal{E}_k(v) \leq \mathcal{E}_0(v)$ , i.e. l'énergie associée à la labellisation erronée est *meilleure* que celle associée à la labellisation de référence. Ces occurrences peuvent être expliquées essentiellement par deux types de phénomènes. D'une part, nous avons observé que les régions de petite taille risquent d'être mal labellisées lorsque leur taux d'occupation au sein d'un pixel basse résolution est trop faible pour apporter une contribution positive à l'énergie. Dans certains cas, un label erroné peut être attribué à une petite région de manière à compenser l'erreur signée due à la contribution des autres régions (correctement labellisées) présentes dans le pixel basse résolution. D'autre part, puisque les distributions gaussiennes des classes que nous considérons se recouvrent partiellement, une occurrence dans la queue de distribution d'une classe donnée risque d'avoir une probabilité supérieure d'appartenir à la classe adjacente et, par conséquent, de se voir attribuer un label incorrect. L'utilisation de données multitemporelles, comme dans le cas réel, pourrait réduire ce type d'erreurs puisque l'information temporelle rend les classes plus séparables.

## 6.6. Sensibilité au facteur de résolution

Les méthodes sous-pixeliques sont souvent très limitées (cf. chapitre 3) par le facteur d'échelle considéré. La méthode que nous proposons a été développée dans l'objectif de traiter de forts rapports de résolution, tels que  $15 \times 15$  pour l'utilisation couplée des capteurs **SPOT/HRVIR** et **MeRIS** ou  $50 \times 50$  pour **SPOT/HRVIR** et **SPOT/VGT**. Dans cette section, nous proposons d'analyser l'évolution des performances de la méthode en fonction du rapport de résolution et de la taille des régions relativement au pixel basse résolution.

Soit  $N = \frac{|\mathcal{D}_{HR}|}{|\mathcal{D}_{BR}|}$  le rapport de résolution entre les images haute et basse résolution considérées. A partir du jeu de 165 images **HR** simulées (cf. section 10.2), les images **BR** ont été simulées pour les rapports de résolution suivants :  $15 \times 15$ ,  $30 \times 30$  et  $50 \times 50$ .

Compte tenu des limites *naturelles* qu'engendrent la présence de "trop" petites régions dans la segmentation, nous calculons le taux d'occupation de toutes les régions mal labellisées. La figure 6.8 (a)-(c) présente l'histogramme du taux d'occupation des régions mal labellisées pour des facteurs de résolution de  $15 \times 15$ ,  $30 \times 30$  et  $50 \times 50$ . Remarquons que ces histogrammes sont très proches, montrant ainsi que la performance de l'algorithme dépend de la taille relative des régions par rapport aux pixels basse résolution mais pas du facteur de résolution en lui-même. De plus, ces histogrammes montrent un pic d'occurrences pour les régions occupant une surface inférieure à 1% du pixel basse résolution. Cette observation pourrait être exploitée pour contraindre la segmentation à la création de régions représentant plus de 1% du pixel basse résolution. Un pré-traitement de ce type conduirait probablement à diminuer de moitié les erreurs de labellisation. Par ailleurs, il semble raisonnable de ne pas s'intéresser à des objets de si petite taille devant l'échelle d'observation.



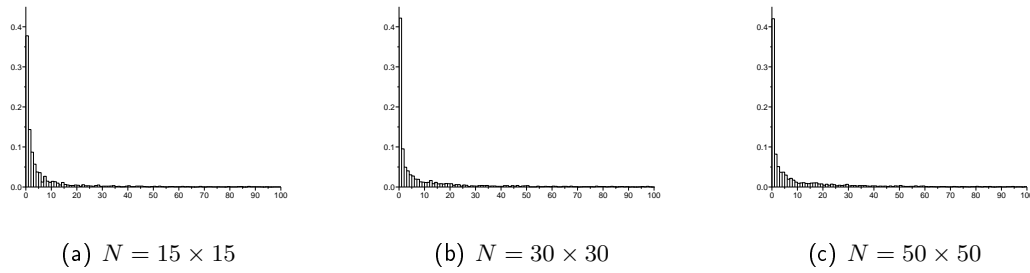


FIG. 6.8.: Histogramme du taux d'occupation des régions mal labellisées au sein des pixels basse résolution pour différents facteurs de résolution. L'axe des abscisses représente le taux d'occupation d'une région mal labellisée au sein d'un pixel basse résolution et l'axe des ordonnées représente le pourcentage d'occurrence.

Le tableau 6.9 présente des statistiques réalisées sur le taux d'occupation des régions mal labellisées dans le pixel basse résolution, selon le facteur de résolution considéré. Remarquons

N	10 ème percentile	médiane	70 ème percentile	80 ème percentile	90 ème percentile
$15^2$	0.44	2	5	10	25.7
$30^2$	0.11	1.77	7.7	16.7	41.4
$50^2$	0.08	2	8.13	17.6	39.8

FIG. 6.9.: Statistiques du taux d'occupation d'une région mal labellisée dans un pixel basse résolution en fonction du facteur de résolution considéré ( $N$ ).

que 50% des régions représentent moins de 2% du pixel basse résolution. Dans la littérature, la plupart des méthodes qui traitent des images **BR** dans le but d'estimer les proportions des classes d'occupation du sol considèrent des parcelles représentant plus de 5% du pixel basse résolution (cf. [Le Hégarat-Masclé et al., 2005]).

En ce qui concerne la performance de la méthode en fonction du facteur de résolution considéré, la figure 6.10 montre que les pourcentages d'erreurs de labellisation sont assez étalés. Cette variabilité peut être attribuée au fait que les expériences ont été réalisées dans un contexte monotemporel alors que les caractéristiques de certaines classes sont parfois très proches. De plus, la ligne médiane montre de bons résultats pour un facteur de résolution allant jusqu'à  $30 \times 30$ , aussi bien pour la version supervisée que pour la version non-supervisée (avec moins de 5% de pixels mal labellisés pour  $N = 30 \times 30$ ). Au-delà, les performances diminuent avec environ 20% de pixels mal labellisés pour un facteur de résolution  $50 \times 50$ . La dégradation des performances lorsque le facteur de résolution augmente est donc manifeste, non-linéaire. Notons, cependant, que le nombre d'expériences réalisées pour la figure 6.10(b) est moins conséquent que celui utilisé pour 6.10(a) en raison du temps de calcul nécessaire à l'algorithme non-supervisé, nettement supérieur à la version

supervisée. Ce graphe est donc difficile à interpréter, en particulier pour le rapport de résolution  $N = 50 \times 50$  où les résultats sont très hétérogènes.

Remarquons que, contrairement au cas étudié figure 6.8 où la taille des régions était considérée en termes de pourcentage d'occupation dans un pixel **BR**, la figure 6.10 présente les pourcentages d'erreurs de labellisation obtenues en fonction du rapport de résolution  $N$ . L'analyse couplée des figures 6.8 et 6.10 permet de mettre en évidence le fait que les performances de la labellisation se dégradent principalement lorsque le rapport de résolution augmente, parce que la proportion d'une région donnée dans un pixel **BR** est d'autant plus faible que ce rapport est grand. Les limites observées figure 6.10 lorsque  $N \geq 30$  ne constituent pas des limites absolues de la méthode. Elles correspondent à un jeu de données simulées à partir d'extraits d'images réelles de la Plaine du Danube, où les parcelles sont de petite taille contrairement, typiquement, à un parcellaire américain où les parcelles sont souvent vastes.

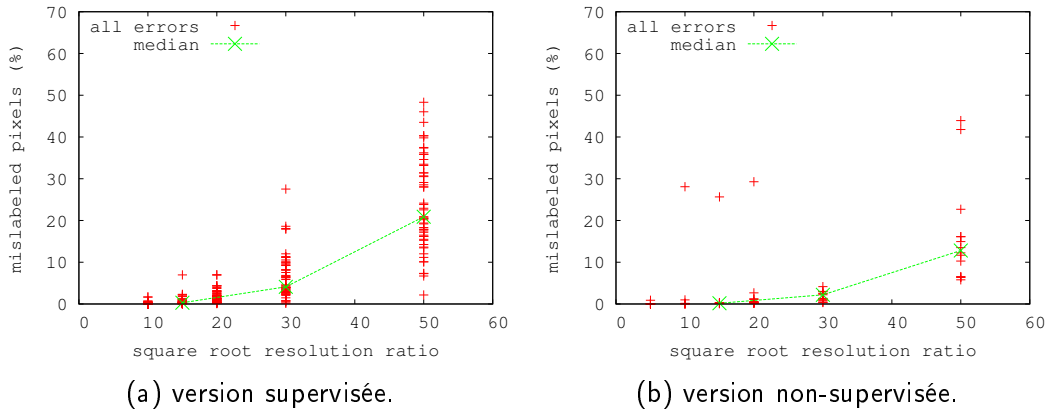


FIG. 6.10.: Chaque croix représente le pourcentage des pixels mal labellisés obtenus pour un jeu de données et un rapport de résolution fixé avec la version supervisée (a) et avec la version non-supervisée (b). Le rapport de résolution  $N$  est représenté en abscisses et le pourcentage en ordonnées. La ligne verte superposée correspond à la médiane de ces valeurs de pourcentage.

Les résultats présentés dans cette section, dans un contexte monotemporel, montrent une certaine robustesse de la méthode vis-à-vis du rapport de résolution. En effet, lorsque ce rapport augmente, les résultats semblent stables en terme de taille des régions relativement à la taille des pixels **BR**. Cependant, pour le jeu de données considéré, la figure 6.10 montre que la segmentation est trop fine pour l'utilisation de données **BR** avec un rapport de résolution  $N = 50 \times 50$ . Pour améliorer ces résultats, la segmentation utilisée pourrait être contrainte selon la grille de pixels **BR**, de manière à ce que toutes les régions de la segmentation occupent plus de 10% environ (d'après la figure 6.8) des pixels **BR**. Les résultats seraient alors optimisés en fonction du rapport de résolution considéré.

Pour finir, nous proposons d'analyser les performances de la méthode en fonction du

rapport de résolution dans un cadre simple d'application aux données réelles. Dans cet objectif, la méthode de labellisation est appliquée à des images pseudo-réelles dans un contexte monotemporel supervisé. Nous utilisons les mêmes extraits de la base de données ADAM que ceux utilisés dans la section 6.3, avant simulation, soit 165 images réelles **HR**. Ces images réelles (**SPOT/HRV**) sont moyennées par blocs de taille  $5 \times 5$  à  $50 \times 50$  pour créer des images **BR** dites pseudo-réelles. La figure 6.11 présente le pourcentage de pixels mal labellisés en fonction du rapport de résolution considéré. Elle montre, comme dans le cas des données simulées, que le pourcentage d'erreurs obtenu présente une certaine variabilité. Cependant, les pourcentages d'erreurs sont significativement plus importants que dans le cas simulé. En effet, la labellisation à partir de données réelles comprend, en médiane, 4.5% d'erreurs pour un rapport de résolution de  $15^2$  alors que, dans le cas de données simulées, seules 0.3% d'erreurs ont été observées (en médiane). Cette différence

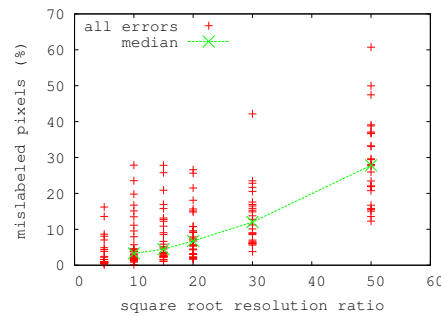


FIG. 6.11.: Pourcentage des pixels mal labellisés obtenus avec l'algorithme supervisé, à partir d'images réelles monotemporelles.

significative peut être attribuée au fait que les images réelles monotemporelles ne sont pas toujours bien modélisées par l'hypothèse gaussienne. De plus, le jeu de données utilisé contient des images acquises en hiver, période où la réponse spectrale des sols nus peut être très spécifique (modélisable par une densité de probabilité particulière).

Ces expériences réalisées sur 165 extraits d'images pseudo-réelles en utilisant la version supervisée de l'algorithme de labellisation montrent que, pour une segmentation donnée, la qualité de la labellisation est très limitée lorsque le rapport de résolution considéré est supérieur à 30. En particulier, pour un rapport  $N = 50 \times 50$ , 30% des pixels sont alors mal labellisés. Ces résultats montrent que la méthode ne permet une labellisation satisfaisante, sur ces données, que si  $N \leq 20 \times 50$ .

Troisième partie .

## Détection de changements à l'échelle sous-pixelique



## 7. Un modèle *a contrario* pour la détection de changements

Dans ce chapitre, nous présentons une méthode de détection de changements basée sur la comparaison entre une séquence d'images **BR** (basse résolution) et une classification **HR** (haute résolution) décrivant un état antérieur de la même scène. Toutes ces images sont supposées géométriquement géoréférencées entre elles, *i.e.* l'ensemble des pixels **HR** correspondants à chaque pixel **BR** est connu. La méthode développée est basée sur l'élaboration d'un critère probabiliste capable d'évaluer le degré de cohérence d'une séquence d'images vis-à-vis d'une classification. Elle ne repose pas sur la modélisation *a priori* des données mais, au contraire, sur le rejet d'un modèle *non structuré*, dit modèle *a contrario*, par l'observation de données structurées.

### 7.1. Modélisation du problème

Notre objectif, dans cette partie, est de détecter l'ensemble des pixels pour lesquels une séquence temporelle d'images **BR** ne correspond plus à la classification **HR** de référence. Cette classification, qui peut être vue comme une labellisation associée à une interprétation de chacun des labels, n'est utilisée ici que pour sa labellisation. En effet, ce qui nous intéresse est davantage le fait que différentes régions aient le même type d'occupation du sol que leur type d'occupation en lui-même. La labellisation, notée  $\lambda$ , décrit la scène observée à une date  $t_0$  donnée, date de référence pour la détection de changements. Formellement, il s'agit d'une application  $\lambda : \mathcal{D}_{\text{HR}} \rightarrow \mathcal{L} = \{1, \dots, L\}$  qui associe à chaque pixel du domaine  $\mathcal{D}_{\text{HR}}$  un label appartenant à  $\mathcal{L}$  qui représente son type d'occupation du sol. Dans le but de détecter les changements qui apparaissent au cours d'une période donnée (séquence d'images **BR**) par rapport à la date de référence (labellisation  $\lambda$ ), nous proposons, dans cette section, de commencer par modéliser le problème considéré. Pour simplifier, nous le présentons dans le cas particulier d'une séquence d'images réduite à une seule image **BR**, notée  $v$ . Le cas général s'en déduit directement en considérant un vecteur d'images **BR**.

Selon la labellisation  $\lambda$ , chaque pixel du domaine  $\mathcal{D}_{\text{HR}}$  est "pur", *i.e.* représente un seul type d'occupation du sol. L'image **BR**, en revanche, est constituée principalement de pixels dits "mixtes" puisque la surface représentée dans un pixel basse résolution est généralement composée de plusieurs parcelles, éventuellement d'occupation du sol différentes. En effet, avec les

capteurs que nous considérons pour cette application, un pixel basse résolution peut couvrir une surface allant de  $300 \times 300 \text{ m}^2$  à  $1 \text{ km}^2$ . Dans ce cas, l'intensité moyenne mesurée en un pixel basse résolution est déterminée par des réponses spectrales des différents types d'occupation du sol qui y sont représentés. Comme dans la partie II, nous supposons que cette intensité moyenne est égale à la somme des valeurs caractéristiques (moyenne, en général) de chaque type d'occupation du sol pondérée par leur taux de représentation au sein du pixel. Cette hypothèse, couramment appelée *modèle linéaire de mélange*, est très souvent utilisée pour le problème de la désaggrégation de pixels mixtes [Faivre et Fischer, 1997], [Faivre et al., 2000], [Cardot et al., 2003], [Settle et Campbell, 1998] mais aussi pour le mélange de plusieurs bandes spectrales [Van Der Meer, 1999], [Zhukov et al., 1999]. Soit  $y$  un pixel appartenant au domaine basse résolution  $\mathcal{D}_{\text{BR}}$  et  $\alpha_l(y)$  la proportion du pixel  $y$  occupée par le label  $l$ . Par définition, pour tout pixel  $y$ , la somme des proportions de chaque label vaut 1, i.e.  $\sum_{l \in \mathcal{L}} \alpha_l(y) = 1$ . Si  $\mu = (\mu(l))_{l \in \mathcal{L}}$  représente l'intensité moyenne caractéristique de la classe de label  $l$ , le modèle linéaire de mélange permet alors d'estimer l'intensité d'un pixel basse résolution  $y$  à partir des proportions de chaque classe et de leur intensité moyenne par

$$\hat{v}(y, \mu) = \sum_{l \in \mathcal{L}} \alpha_l(y) \mu(l). \quad (7.1.1)$$

Dans cette étude, c'est la moyenne statistique correspondant à chaque label qui est considérée comme intensité moyenne caractéristique. Nous définissons à présent la distance entre l'intensité observée  $v(y)$  en un pixel  $y$  et son estimation  $\hat{v}(y, \mu)$ , pour tous les pixels d'un sous-domaine  $D$  de  $\mathcal{D}_{\text{BR}}$ . Supposons que l'image **BR**  $v$  soit une réalisation d'un champ aléatoire  $V$ . Par souci de simplicité, notons  $v_D = (v(y))_{y \in D}$  et  $V_D = (V(y))_{y \in D}$  les restrictions au sous-domaine  $D$ , respectivement, de l'image observée et du champ aléatoire dont elle est une réalisation. Cette distance résiduelle peut être mesurée par la norme euclidienne

$$\delta_\mu(v_D) = \|(v(y) - \hat{v}(y, \mu)) \mathbf{1}_D(y)\|_2, \quad (7.1.2)$$

$$= \sqrt{\sum_{y \in D} (v(y) - \hat{v}(y, \mu))^2}. \quad (7.1.3)$$

Dans le contexte que nous nous sommes fixé, les moyennes caractéristiques des classes  $\mu(l)$  sont inconnues. Elles peuvent, par exemple, être estimées par régression linéaire, i.e. en cherchant la famille  $\mu$  qui minimise  $\delta_\mu(v_D)$ . Soit, dans le cas monodimensionnel,

$$\delta(v_D) = \min_{\mu \in \mathbb{R}^{\mathcal{L}}} \delta_\mu(v_D), \quad (7.1.4)$$

la distance minimale mesurée sur un sous-domaine  $D$  de  $\mathcal{D}_{\text{BR}}$ .

Le choix de la norme euclidienne pour mesurer la distance entre l'intensité observée et l'intensité estimée est naturel dans un contexte de régression linéaire. Cependant, la norme sup a aussi été envisagée mais le fait qu'elle conduise automatiquement à rejeter les événements rares n'incite pas à l'utiliser pour ce type d'applications. En effet, il est généralement facile de diminuer

une distance définie par la norme sup en éliminant quelques échantillons, qui correspondent probablement à la queue de distribution. De plus, les moyennes  $\mu$  étant inconnues en pratique, le choix de cette norme est directement associé à celui de l'estimateur. Les estimateurs définis à partir de la norme sup étant parmi les moins robustes, nous ne considérons pas davantage cette possibilité.

A ce stade, la principale difficulté de la détection est alors la définition d'un seuil *a priori* sur l'erreur résiduelle  $\delta(v_D)$ , seuil au-delà duquel décider de l'apparition de changements sur le sous-domaine  $D \in \mathcal{D}_{BR}$ . Dans la section 7.2, nous construisons un modèle de détection *a contrario* permettant de combiner les différents paramètres en un seul critère probabiliste capable de détecter le sous-domaine sur lequel une image est cohérente avec une labellisation tout en contrôlant le nombre moyen de fausses alarmes attendu. Le complémentaire de ce sous-domaine représente alors l'ensemble des pixels susceptibles de représenter des changements.

## 7.2. Détection *a contrario*

Le principe de détection *a contrario*, introduit en analyse d'images par [Desolneux et al., 2000], permet de calculer un niveau de significativité sans avoir à modéliser les changements, ni à quantifier les écarts attendus (bruit, distortions, variabilité intrinsèque, etc.), *i.e.* en introduisant très peu d'information *a priori*. Cette approche systématique est basée sur un principe simple de perception attribué à Helmholtz. Selon ce principe, notre perception repose sur la détection de structures qui s'opposent au hasard. Il suggère donc de détecter des structures (lignes, courbes, etc.) en tant qu'événement rare par rapport à un modèle naïf, aléatoire (non structuré). Cette idée est essentielle pour la généralité des modèles *a contrario* puisqu'elle permet de détecter des structures sans chercher à les modéliser précisément. Formellement, l'approche *a contrario* repose sur une mesure de la significativité d'une détection par l'espérance du nombre de fausses alarmes plutôt que la probabilité d'une fausse alarme comme pour les approches statistiques classiques [Bonferroni, 1936], [Hommel, 1988].

**Définition 7.2.1** Soit  $E$  l'ensemble des événements potentiellement détectables, et  $\mathbb{P}$  la probabilité sur  $E$  d'apparition des événements. A tout événement  $e \in E$  est associé son nombre de fausses alarmes, défini par

$$NFA(e) = |E| \cdot \mathbb{P}(e) \quad (7.2.5)$$

et pour tout réel  $\varepsilon > 0$ ,  $e$  est dit  $\varepsilon$ -significatif si et seulement si

$$NFA(e) \leq \varepsilon. \quad (7.2.6)$$

Cette définition ne peut être utilisée qu'après avoir défini un modèle naïf suivant le principe de Helmholtz, modèle permettant de donner un sens à la probabilité  $\mathbb{P}(e)$ . Une telle définition permet de garantir que l'espérance du nombre de fausses alarmes soit inférieur à  $\varepsilon$  (*cf.* preuve de la proposition 7.2.4). L'utilisation de l'espérance permet de s'affranchir du fait que les détections multiples ne soient pas indépendantes pour déterminer le domaine de détection de manière absolue.



Il permet de contrôler le nombre moyen de fausses détections et de déterminer automatiquement le seuil de détection en conséquence. La méthode est alors réduite à un seul paramètre : le nombre de fausses alarmes.

Pour fixer les idées sur le principe de cette approche et l'avantage d'utiliser des espérances plutôt que des probabilités, reprenons l'exemple des dates d'anniversaire proposé par [Desolneux et al., 2006]. Dans un groupe de  $N$  personnes, supposons que deux d'entre elles ont la même date d'anniversaire. Cet événement est-il significatif, *i.e.* est-il possible que ce soit arrivé par hasard ou y a-t-il une raison à cela ? En accord avec le principe de Helmholtz, supposons, en négligeant les années bissextiles, que toutes les dates d'anniversaire aient la même probabilité d'occurrence  $1/365$ . Pour évaluer la significativité de l'événement considéré, nous avons deux possibilités :

1. Calculer la probabilité que deux personnes, au moins, aient la même date d'anniversaire :

$$1 - 1 \cdot \left(1 - \frac{1}{365}\right) \cdot \left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{N-1}{365}\right). \quad (7.2.7)$$

Une tabulation est alors nécessaire pour savoir si cette probabilité est inférieure à  $\varepsilon$ .

2. Suivant la définition 7.2.5, considérons l'ensemble  $E$  des événements  $e_{kl} =$  "les personnes  $k$  et  $l$  ont la même date d'anniversaire", avec  $1 \leq k < l \leq N$ , et la mesure  $\mu$  définie par  $\mu(e_{kl}) = 1$  si les dates d'anniversaire des personnes  $k$  et  $l$  sont les mêmes et  $\mu(e_{kl}) = 0$  sinon. Le cardinal de  $E$  vaut  $\frac{N(N-1)}{2}$  et la probabilité  $\mathbb{P}(\mu(e_{kl}) = 1)$  vaut  $\frac{1}{365}$ . Par conséquent, l'évènement  $\mu(e_{kl}) = 1$  est  $\varepsilon$ -significatif si

$$NFA(e_{kl}) = \frac{N(N-1)}{2} \cdot \frac{1}{365} \leq \varepsilon. \quad (7.2.8)$$

Ce critère est beaucoup plus simple que (7.2.7) puisqu'il relie quadratiquement  $N$  à  $\varepsilon$ .

Récemment, [Grosjean et Moisan, 2006] ont proposé la définition suivante du nombre de fausses alarmes dans un cadre plus général.

**Définition 7.2.2** Soit  $(X_i)_{1 \leq i \leq N}$  un ensemble de variables aléatoires. Une fonction  $F(i, x)$  est un *NFA* (nombre de fausses alarmes) pour les variables aléatoires  $(X_i)$  si

$$\forall \varepsilon, \mathbb{E} [|\{i, F(i, X_i) \leq \varepsilon\}|] \leq \varepsilon. \quad (7.2.9)$$

Avec cette définition, si une fonction  $F$  vérifie (7.2.9), alors la famille de tests  $F(i, X_i) \leq \varepsilon$  garantit un nombre moyen de fausses alarmes inférieur à  $\varepsilon$ . Le principe de la détection *a contrario* permet donc une certaine objectivité dans la mesure où il utilise très peu d'information *a priori*. Il permet aussi de réduire tous les paramètres nécessaires à la détection à un seul critère : l'espérance du nombre de fausses alarmes. Ce dernier point est essentiel puisque, en pratique, il rend la détection entièrement automatique en fixant le nombre attendu de fausses alarmes à 1 par exemple. Ce type de détection a déjà été appliqué, notamment, à la détection d'alignements [Desolneux et al., 2000], de bords contrastés [Desolneux et al., 2001], de modes d'histogrammes et d'amas de points [Desolneux et al., 2003], de rigidités entre deux nuages de points [Moisan et Stival, 2004].

Le problème de la détection de changements sur la surface terrestre à partir de données de télédétection doit tenir compte d'un certain nombre de facteurs spécifiques qui rendent inefficaces nombre de méthodes d'analyse d'images multitemporelles utilisées habituellement pour d'autres applications. Par exemple, les variations d'illumination et d'humidité du sol, les différences de calibration du capteur entre les dates considérées, l'absence d'information *a priori* sur les changements attendus, ou les imperfections de recalage sont autant de facteurs à prendre en compte, même s'ils donnent lieu en général à une série de prétraitements spécifiques (corrections radiométriques, recalage, etc.). Les changements susceptibles de se produire sur la surface terrestre tels que, par exemple, les rotations de culture, la contamination de parcelles agricoles, les coupes de forêt ou les incendies, peuvent être de taille et d'intensité radiométrique très variées. Dans l'objectif d'élaborer une méthode générique de détection de changements, il n'est donc pas envisageable de lister tous les changements possibles pour définir un modèle *a priori* des changements. De plus, les profils d'évolution temporelle des différents types d'occupation du sol sont variables d'une année à l'autre, et d'une zone géographique à une autre. Un modèle *a priori* sur les non-changements semble alors aussi difficile à définir physiquement.

Nous proposons de nous placer dans le cadre de la modélisation *a contrario* pour détecter un sous-domaine de l'image en tant que *grande déviation* à partir d'un modèle générique. Etant donnée une labellisation  $\lambda$  (image **HR** constante par morceaux) décrivant l'état de la surface étudiée à une date  $t_0$  fixée, et une image **BR** de la même scène acquise ultérieurement à une date  $t$  ( $t > t_0$ ), nous proposons de détecter dans quel sous-domaine spatial  $D$  la labellisation  $\lambda$  est encore correcte à la date  $t$ . L'ensemble des pixels **BR** correspondant à un changement de type d'occupation du sol est alors le complémentaire de  $D$ . Dans cet objectif, nous introduisons une mesure de *cohérence* entre la labellisation haute résolution et l'image **BR** observée fondée sur le degré de contradiction qu'elle implique en référence à un modèle *non structuré*, dit modèle *a contrario*. La méthode que nous proposons repose donc sur la définition d'un modèle naïf, dit *a contrario*.

**Définition 7.2.3 (modèle *a contrario* pour la BR)** *Le modèle *a contrario* ( $H_0(m)$ ) pour l'image basse résolution est un champ aléatoire  $V$  de  $|\mathcal{D}_{BR}|$  variables gaussiennes i.i.d.  $\mathcal{N}(m, \sigma^2 I_{|\mathcal{D}_{BR}|})$  où  $m \in \mathbb{R}^{|\mathcal{D}_{BR}|}$  et  $\sigma > 0$  sont fixés.*

Le choix des paramètres du modèle naïf est discuté dans la section 7.3. L'écart-type  $\sigma$  est considéré comme fixé, et nous étudierons plusieurs possibilités pour le vecteur des moyennes. Avant de discuter les valeurs de ces paramètres, poursuivons la description du modèle de détection dans le cas général.

Suivant le principe de Helmholtz, décrit pour la première fois en analyse d'images par [Lowe, 1985], deux éléments peuvent être groupés si leur position a une probabilité très faible de résulter d'un arrangement accidentel ; nous nous intéressons aux sous-domaines de l'image pour lesquels l'erreur quadratique mesurée entre l'image et son estimée est trop petite pour être raisonnablement expliquée par le hasard. En fait, il s'agit de *s'étonner* que l'intensité observée sur

un sous-domaine de  $\mathcal{D}_{BR}$  soit particulièrement proche, pour une certaine famille de valeurs  $\mu$ , de l'intensité estimée à partir des proportions de chaque label et de la famille  $\mu$ . Une des principales difficultés est alors la définition objective et automatique d'un seuil *a priori* à partir duquel l'erreur  $\delta(v_D)$  n'est plus considérée comme acceptable. Un seuil de détection  $\delta_D$  pourrait être choisi de manière à garantir que

$$\mathbb{P}(\exists D, \delta(V_D) \geq \delta_D) \leq \varepsilon, \quad (7.2.10)$$

où  $\varepsilon$  est un paramètre fixé (par exemple  $10^{-3}$ ) et  $\delta(V_D)$  l'erreur quadratique obtenue en considérant le champ aléatoire  $V$  décrit par le modèle *a contrario* (image de bruit blanc gaussien). Le paramètre  $\varepsilon$  permet de contrôler la fiabilité du test. En effet, plus  $\varepsilon$  est petit, plus le test est exigeant et fiable. Cependant, les dépendances entre les variables aléatoires  $(\delta(V_D))_{D \in \mathcal{D}_{BR}}$  sont très difficiles à estimer, ce qui rend impossible le calcul explicite de  $\mathbb{P}(\exists D, \delta(V_D) \geq \delta_D)$ .

Suivant les travaux [Desolneux et al., 2000, Desolneux et al., 2001, Desolneux et al., 2003], nous proposons de mesurer l'espérance du nombre de fausses détections plutôt que de contrôler la probabilité de fausses détections. Le critère que nous proposons permet d'établir une mesure de cohérence pour un sous-domaine vis-à-vis d'une labellisation. Il est modélisé de manière à garantir que le nombre de détections obtenues sur des données aléatoires soit aussi faible que souhaité. Pour cela, une quantification du nombre de tests (nombre de sous-domaines étudiés) est nécessaire.

Le nombre de tests  $\eta$  est un coefficient de pondération capable de contrôler le nombre moyen de fausses détections et, par la même occasion, d'éviter l'introduction d'un seuil de détection. Cette définition est guidée non seulement par la contrainte de ne pas détecter à tort, mais aussi par celle d'obtenir une méthode à un paramètre unique : le nombre moyen de fausses alarmes acceptable. En pratique, il suffit de le fixer à 1 pour obtenir une méthode automatique. Les méthodes statistiques classiques sont aussi, typiquement, dépendantes d'un seul paramètre mais il s'agit en général du seuil de décision sur la probabilité maximale. Ce type de seuil est très dépendant du contexte et des images utilisées. Dans l'objectif de contrôler le nombre de fausses alarmes, [Desolneux et al., 2000] ont introduit la notion d'événements  $\varepsilon$ -significatifs permettant de déterminer automatiquement le seuil de détection en fonction d'un unique paramètre : le nombre maximal de fausses détections autorisé.

**Proposition 7.2.4 (nombre de fausses alarmes)** Soient  $\eta : \mathbb{N} \rightarrow \mathbb{N}$  et, pour tout sous-domaine  $D \subset \mathcal{D}_{BR}$ ,  $\mathbb{P}_{H_0}[\delta(V_D) \geq \delta(v_D)]$  la probabilité d'observer une erreur minimale inférieure à  $\delta(V_D)$  sous l'hypothèse  $H_0(m)$ . La relation

$$NFA(D, \delta(v_D), \sigma, m) = \eta(|D|) \cdot \mathbb{P}_{H_0}[\delta(V_D) \geq \delta(v_D)] \quad (7.2.11)$$

définit un nombre de fausses alarmes dès que

$$\sum_{D \in \mathcal{D}_{BR}} \frac{1}{\eta(|D|)} \leq 1.$$

Soit  $\varepsilon > 0$ , un sous-domaine  $D$  de  $\mathcal{D}_{BR}$  est dit  $\varepsilon$ -significatif si  $NFA(D, \delta(v_D), \sigma, m) \leq \varepsilon$ .

**Démonstration.**

D'après la définition 7.2.2,

$$NFA(D, \delta(v_D), \sigma, m) = \eta(|D|) \cdot \mathbb{P}_{H_0}[\delta(V_D) \geq \delta(v_D)] \quad (7.2.12)$$

définit un nombre de fausses alarmes si, pour tout  $\varepsilon > 0$ , l'espérance du nombre de domaines  $\varepsilon$ -significatifs, sous  $H_0(m)$ , est inférieure à  $\varepsilon$ . Notons  $\chi_D$  la variable aléatoire binaire qui vaut 1 si le sous-domaine  $D$  de  $\mathcal{D}_{BR}$  est  $\varepsilon$ -significatif, 0 sinon, *i.e.*

$$\chi_D = \begin{cases} 1 & \text{si } NFA(D, \delta(v_D), \sigma, m) \leq \varepsilon \\ 0 & \text{si } NFA(D, \delta(v_D), \sigma, m) > \varepsilon. \end{cases}$$

La variable aléatoire  $NDS_\varepsilon = \sum_{D \in \mathcal{D}_{BR}} \chi_D$  représente alors le nombre de sous-domaines  $\varepsilon$ -significatifs dans  $\mathcal{D}_{BR}$ . Par linéarité, l'espérance du nombre de sous-domaines  $\varepsilon$ -significatifs est

$$\begin{aligned} \mathbb{E}[NDS_\varepsilon] &= \sum_{D \in \mathcal{D}_{BR}} \mathbb{E}[\chi_D] \\ &= \sum_{D \in \mathcal{D}_{BR}} \mathbb{P}(\chi_D = 1) \end{aligned}$$

Soit  $I$  un intervalle fermé de  $\mathbb{R}$  tel que pour tout sous-domaine  $D$ ,  $\delta(v_D) \in I$ . Pour tout  $D$  sous-domaine de  $\mathcal{D}_{BR}$ , notons  $\delta_\varepsilon^*(D)$  le résidu minimal sur  $D$  tel que le sous-domaine  $D$  soit  $\varepsilon$ -significatif :

$$\delta_\varepsilon^*(D) = \min \left\{ \delta \in I, \mathbb{P}_{H_0}(\delta(V_D) \geq \delta) \leq \frac{\varepsilon}{\eta(|D|)} \right\}. \quad (7.2.13)$$

Montrons que cette valeur est bien définie. Pour cela, considérons la fonction  $f : I \rightarrow [0, 1]$  définie pour tout  $\delta \in I$  par  $f(\delta) = \mathbb{P}_{H_0}[\delta(V_D) \geq \delta]$ . La variable aléatoire  $\delta(V_D)$  est une variable à densité puisqu'elle est définie par

$$\delta^2(V_D) = \min_{\mu \in \mathbb{R}^L} \sum_{y \in D} (V(y) - \hat{v}(y, \mu))^2, \quad (7.2.14)$$

où, pour tout  $y$ , la variable aléatoire  $V(y)$  est à densité et  $\hat{v}(y, \mu)$  est déterministe (*cf.* (7.1.1)). Donc  $(V(y) - \hat{v}(y, \mu))^2$  est une variable aléatoire à densité ainsi que  $\delta^2(V_D)$  en tant que somme de variables aléatoires indépendantes à densité. La fonction de répartition d'une variable à densité étant continue, la fonction  $f$  est continue de  $I$  dans  $[0, 1]$ . Par conséquent,

$$\left\{ \delta \in I, \mathbb{P}_{H_0}(\delta(V_D) \geq \delta) \leq \frac{\varepsilon}{\eta(|D|)} \right\} = f^{-1} \left( \left[ 0, \frac{\varepsilon}{\eta(|D|)} \right] \right), \quad (7.2.15)$$

où  $f^{-1}$  est continue en tant que fonction réciproque d'une fonction continue et monotone (la fonction  $f : \delta \mapsto \mathbb{P}_{H_0}[\delta(V_D) \geq \delta]$  est décroissante). Ainsi,  $\left\{ \delta \in I, \mathbb{P}_{H_0}(\delta(V_D) \geq \delta) \leq \frac{\varepsilon}{\eta(|D|)} \right\}$  est un ensemble fermé et borné en tant qu'image réciproque d'un intervalle fermé et borné par une fonction continue et sa borne inférieure est atteinte.

Donc l'espérance du nombre de domaines  $\varepsilon$ -significatifs satisfait

$$\sum_{D \in \mathcal{D}_{BR}} \mathbb{P}(\chi_D = 1) = \sum_{D \in \mathcal{D}_{BR}} \mathbb{P}_{H_0}(\delta(V_D) \geq \delta_\varepsilon^*(D)) \leq \sum_{D \in \mathcal{D}_{BR}} \frac{\varepsilon}{\eta(|D|)}. \quad (7.2.16)$$

Donc

$$\mathbb{E}[NDS_\varepsilon] \leq \sum_{D \in \mathcal{D}_{BR}} \frac{\varepsilon}{\eta(|D|)}, \quad (7.2.17)$$

et

$$\mathbb{E}[NDS_\varepsilon] \leq \varepsilon \quad (7.2.18)$$

dès que

$$\sum_{D \subset \mathcal{D}_{BR}} \frac{1}{\eta(|D|)} \leq 1. \quad (7.2.19)$$

Il suffit donc de choisir  $\eta$  vérifiant cette inégalité pour garantir un nombre moyen de fausses alarmes inférieur à  $\varepsilon$ .  $\square$

Le choix le plus naturel pour le nombre de tests est de considérer l'ensemble de tous les sous-domaines de  $\mathcal{D}_{BR}$ , soit

$$\eta(|D|) = |\{D, D \subset \mathcal{D}_{BR}\}| = 2^{|\mathcal{D}_{BR}|}. \quad (7.2.20)$$

Remarquons que  $\sum_{D \subset \mathcal{D}_{BR}} \frac{1}{2^{|\mathcal{D}_{BR}|}} = 1$ , donc ce nombre de tests garantit, en moyenne, au plus  $\varepsilon$  fausses détections. De plus, il permet de répartir le risque uniformément sur tous les sous-domaines de  $\mathcal{D}_{BR}$  (quelque soit leur taille et même s'ils se recouvrent). En effet, pour tout sous-domaine  $D$  de  $\mathcal{D}_{BR}$ , la probabilité de détecter  $D$  par erreur (*i.e.* que  $D$  soit  $\varepsilon$ -significatif sous l'hypothèse  $H_0$ ) vaut alors

$$\mathbb{P}_{H_0}(\delta(V_D) \geq \delta(v_D)) = \frac{\varepsilon}{2^{|\mathcal{D}_{BR}|}}, \quad (7.2.21)$$

où le membre de droite ne dépend pas de  $D$ . Un sous-domaine de cardinal  $k$  est alors pénalisé de la même manière qu'un sous-domaine de cardinal  $k' \neq k$ . Par hasard, les domaines de tailles moyennes vont donc être davantage détectés par hasard (*i.e.* sous  $H_0$ ) que ceux de tailles extrêmes puisqu'ils sont plus nombreux. Avec ce nombre de tests, le même poids est attribué à tous les sous-domaines alors qu'intuitivement, pour l'application considérée, on souhaiterait bénéficier d'une estimation plus robuste sur les grands sous-domaines que sur les petits. De plus, il correspond au calcul du nombre total de sous-domaines de l'image, de nombreux recouvrements des sous-domaines sont alors pris en compte.

Le répartition des tests est finalement choisie avec une certaine liberté, elle peut être vue comme une fonction de poids permettant d'exprimer un *a priori* pour la détection. En ce qui concerne l'application visée, nous sommes tentés de favoriser la détection des sous-domaines de grande taille. Nous proposons donc une alternative en déterminant un nombre de tests qui tienne compte de la taille des sous-domaines considérés et qui soit inférieur à  $2^{|\mathcal{D}_{BR}|}$ . Pour cela, nous considérons une répartition du risque capable de conserver un nombre de domaines  $\varepsilon$ -significatifs uniforme sur la taille du sous-domaine : pour tout  $D$ , sous-domaine de  $\mathcal{D}_{BR}$ ,

$$\begin{aligned} \mathbb{P}_{H_0}(\delta(V_D) \geq \delta_\varepsilon^*(D)) &= \frac{\varepsilon}{|\mathcal{D}_{BR}| \cdot |\{D'; |D'| = |D|\}|} \\ &= \frac{\varepsilon}{|\mathcal{D}_{BR}| \cdot C_{|\mathcal{D}_{BR}|}^{|D|}}. \end{aligned}$$

Il suffit alors de considérer le nombre de tests

$$\eta(|D|) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|}, \quad (7.2.22)$$

qui vérifie

$$\begin{aligned} \sum_{D \in \mathcal{D}_{BR}} \frac{1}{\eta(|D|)} &= \sum_{D \in \mathcal{D}_{BR}} \frac{1}{|\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|}} \\ &= \sum_{k=1}^{|\mathcal{D}_{BR}|} \sum_{|D|=k} \frac{1}{|\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|}} \\ &= \sum_{k=1}^{|\mathcal{D}_{BR}|} C_{|\mathcal{D}_{BR}|}^{|D|} \frac{1}{|\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|}} \\ &= |\mathcal{D}_{BR}| \frac{1}{|\mathcal{D}_{BR}|} \\ &= 1. \end{aligned}$$

Ce choix satisfait donc la proposition 7.2.4. Il garantit un nombre moyen de domaines  $\varepsilon$ -significatifs inférieur à  $\varepsilon$  tout en répartissant le risque en fonction du cardinal des sous-domaines considérés. Cette répartition, symétrique, traite de la même façon les domaines de tailles extrêmes (petits et grands) de l'image. Si elle permet de ré-équilibrer le risque en faveur des domaines de grande taille, elle agit de même sur les domaines de petite taille ce qui n'est pas particulièrement recherché. Cependant, en pratique, les changements ont tendance à affecter plutôt une partie minoritaire de l'image ( $< 50\%$ ) donc c'est ce nombre de tests que nous considérons dans la suite de l'étude.

Le paramètre  $\varepsilon$  permet de fixer le nombre moyen de fausses détections autorisé. En pratique, nous choisissons de poser  $\varepsilon = 1$  de sorte à garantir, en moyenne, une fausse détection. Par ailleurs, une fois ce paramètre fixé, la méthode est alors entièrement automatique.

Ainsi défini, plus le *NFA* est faible, plus le sous-domaine  $D$  est cohérent avec le modèle de l'image. Plusieurs stratégies sont alors envisageables. Par exemple, nous pourrions décider de rechercher le sous-domaine  $D$  de taille maximale et tel que le nombre de fausses alarmes qui lui est associé soit inférieur à un seuil  $NFA_{min}$ . Cependant, nous pouvons considérer que la formulation du *nombre de fausses alarmes* adoptée prend déjà en compte la taille du domaine étudié et, pour des questions de robustesse, nous préférons retenir le domaine le plus significatif. Nous proposons, par conséquent, de rechercher le sous-domaine  $D$  qui minimise le *nombre de fausses alarmes*. Avec l'hypothèse *a contrario* donnée par la définition 7.2.3, le *NFA* peut être calculé explicitement.

**Théorème 7.2.5** Avec l'hypothèse *a contrario* donnée par  $H_0(m)$ , quelque soit  $m \in \mathbb{R}^{|\mathcal{D}_{BR}|}$ , le nombre de fausses alarmes associé à un sous-domaine  $D$  d'une image  $v$  est déterminé par

$$NFA(D, \delta(v_D), \sigma, m) = \eta(|D|) \cdot f(|D| - |\mathcal{L}|, \delta(v_D), \sigma, m), \quad (7.2.23)$$

où  $\eta(|D|) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|}$  représente le nombre de tests,  $|\mathcal{L}|$  le nombre de labels et, pour tout  $q = |D| - |\mathcal{L}|$ ,

$$f(q, \delta, \sigma, m) = \frac{1}{\sigma^q (2\pi)^{\frac{q}{2}}} \int_{(x_1, \dots, x_q) \in \mathcal{B}_q(\delta)} e^{-\frac{1}{2} \sum_{i=1}^q \left( \frac{x_i - m_i}{\sigma} \right)^2} dx_1 \dots dx_q,$$

où  $\mathcal{B}_q(\delta)$  représente la boule de centre 0 et de rayon  $\delta$  dans  $\mathbb{R}^q$ .

### Démonstration.

Sous l'hypothèse  $H_0$ , le nombre de fausses alarmes est défini (cf. proposition 7.2.4) par

$$NFA(D, \delta(v_D), \sigma, m) = \eta(|D|) \cdot \mathbb{P}_{H_0}[\delta(V_D) \geq \delta(v_D)]. \quad (7.2.24)$$

Pour déterminer explicitement le nombre de fausses alarmes, il faut, étant donné  $\delta(v_D)$ , calculer la probabilité

$$\mathbb{P}_{H_0}[\delta(V_D) \geq \delta(v_D)]. \quad (7.2.25)$$

L'erreur  $\delta$  mesurée sur un sous-domaine  $D$  dépend de la labellisation à travers les proportions  $\alpha_l(y)$  de chaque label  $l$  dans un pixel  $y$  de  $D$ . Soit  $A_D$  la matrice des proportions extraite sur un sous-domaine  $D$ , elle est définie par  $A_D = (\alpha_l(y))_{y \in D}$ . Remarquons que l'erreur résiduelle

$$\delta(V_D) = \min_{\mu \in \mathbb{R}^{\mathcal{L}}} \sqrt{\sum_{y \in D} \left( V(y) - \sum_{l \in \mathcal{L}} \alpha_l(y) \mu(l) \right)^2} \quad (7.2.26)$$

peut être interprétée en termes de distance du vecteur  $V_D$  à l'espace image de la matrice des proportions  $A_D$ , noté  $\text{Im}A_D$ . La distance minimale est alors la norme de la projection orthogonale de  $V_D$  sur l'espace  $(\text{Im}A_D)^\perp$ , i.e.

$$\delta(V_D) = \|p_{(\text{Im}A_D)^\perp}(V_D)\|. \quad (7.2.27)$$

Soit  $n = |D|$ , d'après l'hypothèse *a contrario*, le vecteur  $V_D$  suit une loi normale  $\mathcal{N}(m_D, \sigma^2 I_n)$ , où  $m_D$  représente la restriction du vecteur  $m$ , vu comme une application, au sous-domaine  $D$ . Soit  $P$  la matrice de projection orthogonale de  $p_{(\text{Im}A_D)^\perp}$  alors

$$\begin{aligned} \mathbb{P}[\delta(V_D) \geq \delta(v_D)] &= \mathbb{P}[\|p_{(\text{Im}A_D)^\perp}(V_D)\| \geq \delta(v_D)] \\ &= \mathbb{P}[\|PV_D\| \geq \delta(v_D)]. \end{aligned} \quad (7.2.28)$$

Comme  $\mathbb{R}^n = \text{Im}A_D \oplus (\text{Im}A_D)^\perp$ , il existe une base  $e = (e_1, \dots, e_n)$  dans laquelle les  $q$  premiers vecteurs forment une base de  $\text{Im}A_D$  et les  $n - q$  suivants une base de  $(\text{Im}A_D)^\perp$ . Dans cette nouvelle base, la matrice de projection  $P$  est de la forme

$$\text{Mat}_e(p_{(\text{Im}A_D)^\perp}) = \begin{pmatrix} I_q & 0 \\ 0 & 0_{n-q} \end{pmatrix}$$

où  $q = \dim(\text{Im}A_D)$  et  $n - q = \dim(\text{Im}A_D)^\perp$ . Dans le contexte que nous nous sommes fixé,  $\dim(\text{Im}A_D) = |\mathcal{L}| = n - q$ , i.e.  $q = |D| - |\mathcal{L}|$ .

La probabilité pour que la variable aléatoire  $PV_D$  soit dans une boule  $\mathcal{B}^q(\delta)$  de  $(\text{Im}A_D)^\perp \sim \mathbb{R}^q$ , de rayon  $\delta$ , vérifie

$$\mathbb{P}(PV_D \in \mathcal{B}^q(\delta)) = \mathbb{P}(V_D \in P^{-1}(\mathcal{B}^q(\delta))). \quad (7.2.29)$$

Or  $P^{-1}(\mathcal{B}^q(\delta)) = \mathcal{B}^q(\delta) \times \mathbb{R}^{n-q}$ , et, d'après le modèle *a contrario*, le vecteur aléatoire  $V_D$  suit une loi normale de moyenne  $m_D \in \mathbb{R}^{|D|}$  et de variance  $\sigma^2 I_{|D|}$ . Donc

$$\begin{aligned} & \mathbb{P}(V_D \in P^{-1}(\mathcal{B}^q(\delta))) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\sigma^2 I_n)}} \int_{(x_1, \dots, x_q) \in \mathcal{B}^q(\delta)} \int_{(x_{q+1}, \dots, x_n) \in \mathbb{R}^{n-q}} e^{-\frac{1}{2} \sum_{i=1}^n (\frac{x_i - m_i}{\sigma})^2} dx_1 \dots dx_n \\ &= \frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} \int_{(x_1, \dots, x_q) \in \mathcal{B}^q(\delta)} e^{-\frac{1}{2} \sum_{i=1}^q (\frac{x_i - m_i}{\sigma})^2} dx_1 \dots dx_q \int_{(x_{q+1}, \dots, x_n) \in \mathbb{R}^{n-q}} e^{-\frac{1}{2} \sum_{i=q+1}^n (\frac{x_i - m_i}{\sigma})^2} dx_{q+1} \dots dx_n. \end{aligned}$$

Remarquons que

$$\frac{1}{\sigma^n (2\pi)^{\frac{n}{2}}} = \frac{1}{\sigma^q (2\pi)^{\frac{q}{2}}} \times \frac{1}{\sigma^{n-q} (2\pi)^{\frac{n-q}{2}}}$$

et l'intégrale de la gaussienne en dimension  $n - q$

$$\frac{1}{\sigma^{n-q} (2\pi)^{\frac{n-q}{2}}} \int_{(x_{q+1}, \dots, x_n) \in \mathbb{R}^{n-q}} e^{-\frac{1}{2} \sum_{i=q+1}^n (\frac{x_i - m_i}{\sigma})^2} dx_{q+1} \dots dx_n = 1.$$

Donc, finalement,

$$\mathbb{P}(PV_D \in \mathcal{B}^q(\delta)) = \frac{1}{\sigma^q (2\pi)^{\frac{q}{2}}} \int_{(x_1, \dots, x_q) \in \mathcal{B}^q(\delta)} e^{-\frac{1}{2} \sum_{i=1}^q (\frac{x_i - m_i}{\sigma})^2} dx_1 \dots dx_q. \quad (7.2.30)$$

Le vecteur aléatoire  $PV_D$  est donc gaussien de moyenne  $(m_1, \dots, m_q)$  et de covariance  $\sigma^2 I_q$  sur l'espace  $\text{Im} P$ , et la probabilité  $\mathbb{P}_{H_0}[\delta(V_D) \geq \delta(v_D)]$  est déterminée par la fonction

$$f(q, \delta, \sigma, m) = \frac{1}{\sigma^q (2\pi)^{\frac{q}{2}}} \int_{\mathcal{B}^q(\delta)} e^{-\frac{1}{2} \sum_{i=1}^q (\frac{x_i - m_i}{\sigma})^2} dx_1 \dots dx_q. \quad (7.2.31)$$

□

La distribution obtenue est celle du  $\chi^2$  à  $q$  degrés de liberté. Rappelons qu'une variable  $\chi^2$  correspond à la somme des carrés de variables aléatoires normales centrées réduites. Avec l'hypothèse *a contrario* sur les variables  $(V(y))_{y \in D}$ , la variable aléatoire

$$\delta^2(V_D) = \sum_{y \in D} (V(y) - \hat{V}(y, \mu))^2$$

correspond à la somme de  $|D|$  variables aléatoires normales centrées mais non-réduites, elle ne correspond donc pas directement à un  $\chi^2$ . En revanche, la variable  $\frac{\delta^2(V_D)}{\sigma^2}$  est un  $\chi^2$  donc la probabilité à laquelle nous nous intéressons (7.2.25) vérifie

$$\begin{aligned} \mathbb{P}(\delta^2(V_D) \geq \delta^2) &= \mathbb{P}(\sigma^2 \chi^2 \geq \delta^2) \\ &= \mathbb{P}(\chi^2 \geq \frac{\delta^2}{\sigma^2}). \end{aligned} \quad (7.2.32)$$

La loi du  $\chi^2$  est utilisée pour de nombreuses applications, aussi bien dans le cadre des tests d'adéquation d'observations à une distribution théorique que pour les tests d'indépendance entre deux caractères qualitatifs ou les tests d'homogénéité qui permettent de tester si deux jeux de



variables aléatoires suivent une même loi. Le problème de la détection de changements pourrait être envisagé, par exemple, sous la forme d'un test d'adéquation (cf. Section 7.5).

Avec cette nouvelle mesure, notre objectif est de sélectionner le domaine  $D$  le plus *cohérent* avec la labellisation initiale en tant que domaine minimisant le  $NFA$ . Avant d'étendre le modèle de détection au cas multitemporel, nous discutons section 7.3 le choix du modèle *a contrario*  $H_0$  et de ses paramètres. De plus, le  $NFA$  n'est pas calculable numériquement sous la forme (7.2.31) avec suffisamment de précision. Par conséquent, nous proposons section 7.3 une autre expression du  $NFA$  permettant un calcul numérique plus précis.

### 7.3. Choix du modèle *a contrario*

Le modèle de détection *a contrario* est un modèle générique de détection de cohérence par rapport à un modèle *a contrario*, dit naïf. Ce type de modélisation ne vise pas à décrire au mieux les données, mais simplement à les comparer à un modèle simpliste, *non structuré*, de l'image. Dans la section 7.2, nous avons adopté le modèle  $H_0$  suivant lequel l'image  $\mathbf{BR}$  est un champ aléatoire de variables gaussiennes i.i.d. de moyenne  $m$  et de variance  $\sigma^2$ . Le principe de la méthode est alors de détecter un sous-domaine  $D$  si la prédiction de  $v$  sur  $D$  donnée par l'équation (7.1.1) est trop proche de l'observation pour être raisonnablement expliquée par le hasard.

Le modèle  $H_0$  dépend de deux paramètres : le vecteur des moyennes  $m$  et la matrice de covariance  $\sigma^2 I$ . Dans cette section, nous proposons deux choix de paramètres pour ce modèle ainsi qu'une alternative au modèle  $H_0$ . Pour commencer, nous suggérons d'adopter pour hypothèse naïve le modèle *a contrario* selon laquelle l'image  $\mathbf{BR}$  est de moyenne constante.

**Hypothèse  $H_0^a$**  : le vecteur  $m$  des moyennes du modèle *a contrario*  $H_0(m)$  est de la forme  $\theta(1, 1, \dots, 1)$ , où  $\theta \in \mathbb{R}$  et l'écart-type  $\sigma$  est fixé *a priori*.

**Corollaire 7.3.1** *Sous l'hypothèse  $H_0^a$ , le nombre de fausses alarmes vaut*

$$NFA_a(D, \delta(v_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \cdot f(|D| - |\mathcal{L}|, \delta(v_D), \sigma, 0). \quad (7.3.33)$$

#### Démonstration.

Soit  $A_D$  la matrice des proportions  $(\alpha_l(y))_{y \in D}$ , de taille  $|D| \times |\mathcal{L}|$ . Cette matrice représente, pour chaque ligne  $i$ , les proportions de chaque label contenu dans le pixel  $i$ . La somme en ligne des éléments de la matrice  $A_D$  vaut alors 1, ce qui signifie que l'ensemble des vecteurs proportionnels à  $(1, 1, \dots, 1)$  appartient à l'espace image de  $A_D$ ,  $\text{Im} A_D$ . En particulier, si le vecteur moyenne  $m$  est un vecteur proportionnel à  $(1, 1, \dots, 1)$ , il appartient à l'espace  $\text{Im} A_D$  et satisfait la relation  $A_D m = m$ . Le vecteur moyenne projeté par  $P$  est alors réduit au vecteur nul ( $Pm = 0$ ). D'après la démonstration du théorème 7.2.5, on a  $\mathbb{P}_{H_0}[\delta(V_D) \geq \delta(v_D)] = \mathbb{P}(PV_D \in \mathcal{B}^q(\delta))$  et, d'après le modèle *a contrario* (définition 7.2.3), la variable aléatoire  $PV_D$  suit une loi normale de moyenne  $(m_1, \dots, m_q)$  et de matrice de covariance  $\sigma^2 I_q$ . Sous l'hypothèse  $H_0^a$ ,  $(m_1, \dots, m_q) = (0, \dots, 0)$

et le nombre de fausses alarmes s'écrit

$$NFA_a(D, \delta(v_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \cdot f(|D| - |\mathcal{L}|, \delta(v_D), \sigma, 0).$$

□

Une autre hypothèse possible pour les paramètres du modèle  $H_0$  est de supposer le vecteur des moyennes quelconque et la variance fixée, ce qui revient à considérer une image **BR** dont la valeur moyenne en chaque pixel est quelconque.

**Hypothèse  $H_0^b$**  : le vecteur des moyennes  $m$  du modèle *a contrario*  $H_0$  est un vecteur quelconque de  $\mathbb{R}^{|\mathcal{D}_{BR}|}$  et l'écart-type  $\sigma$  est fixé *a priori*. Comme  $m$  est une inconnue du modèle  $NFA$ , nous cherchons celui qui conduit au *meilleur* modèle dans le sens où il permet de contrôler le *nombre de fausses alarmes* pour tous les modèles  $H_0(m)$ , où  $m \in \mathbb{R}^{|\mathcal{D}|}$ . Pour cela, il suffit de considérer les moyennes  $m$  qui maximisent le *nombre de fausses alarmes*, soit

$$\arg \max_{m \in \mathbb{R}^{|\mathcal{D}|}} NFA(D, \delta(v_D), \sigma, m).$$

Montrons que ce maximum est atteint lorsque la moyenne du modèle  $H_0$  est nulle, ce qui conduit au résultat suivant.

**Proposition 7.3.2** *Sous les hypothèses  $H_0^a$  ou  $H_0^b$ ,*

$$NFA(|D|, \delta(v_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \bar{f}(|D| - |\mathcal{L}|, \delta(v_D), \sigma), \quad (7.3.34)$$

où,

$$\begin{aligned} \bar{f}(q, \delta, \sigma) &\equiv f(q, \delta, \sigma, 0) \\ &= \frac{1}{\sigma^q (2\pi)^{\frac{q}{2}}} \int_{(x_1, \dots, x_q) \in \mathcal{B}_q(\delta)} e^{-\frac{1}{2} \sum_{i=1}^q (\frac{x_i}{\sigma})^2} dx_1 \dots dx_q. \end{aligned} \quad (7.3.35)$$

**Démonstration.**

**Lemme 7.3.3** *Soit  $t > 0$  fixé. La fonction définie pour tout vecteur  $y \in \mathbb{R}^q$  par*

$$g(y) = \int_{\|x\| \leq t, x \in \mathbb{R}^q} e^{-\|x-y\|^2} dx$$

*est maximale quand  $y = 0$ .*

**Preuve du lemme :**

Soient  $x_i$  et  $y_i$  respectivement les  $i$ èmes éléments des vecteurs  $x$  et  $y$ .

$$\begin{aligned} \frac{\partial g}{\partial y_i} &= \int_{\|x\| \leq t, x \in \mathbb{R}^q} -2(y_i - x_i) e^{-\sum_{i=1}^q (y_i - x_i)^2} dx \\ &= \int_{\sum_{j \neq i} x_j^2 < t^2} e^{-\sum_{j \neq i} (y_j - x_j)^2} \left[ \int_{x_i^2 < t^2 - \sum_{j \neq i} x_j^2} -2(y_i - x_i) e^{-(y_i - x_i)^2} dx_i \right] dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_q \end{aligned}$$

En posant  $a(x) = \sqrt{t^2 - \sum_{j \neq i} x_j^2}$ , on obtient

$$\begin{aligned} \int_{x_i^2 < a(x)^2} -2(y_i - x_i) e^{-(y_i - x_i)^2} dx_i &= [e^{-(y_i - x_i)^2}]_{-a(x)}^{a(x)} \\ &= e^{-(y_i - a(x))^2} - e^{-(y_i + a(x))^2}, \end{aligned}$$

et alors

$$\frac{\partial g}{\partial y_i} = \int_{\sum_{j \neq i} x_j^2 < t^2} e^{-\sum_{j \neq i} (y_i - x_i)^2} [e^{-(y_i - a(x))^2} - e^{-(y_i + a(x))^2}] dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_q.$$

La fonction  $g$  est positive et nulle à l'infini donc, par compacité, elle atteint son maximum. De plus, elle est  $\mathcal{C}^\infty$ , donc ce maximum correspond à un point critique.

$$\begin{aligned} \frac{\partial g}{\partial y_i} &= 0 \\ \Leftrightarrow \forall x, e^{-(y_i - a(x))^2} - e^{-(y_i + a(x))^2} &= 0 \\ \Leftrightarrow \forall x, (y_i - a(x))^2 &= (y_i + a(x))^2 \\ \Rightarrow y_i &= 0. \end{aligned}$$

L'unique point critique de la fonction  $g$  est donc  $y = 0$ . Or le maximum global de  $g$  sur  $\mathbb{R}^q$  est un point critique, donc c'est  $y = 0$ .

L'application de ce lemme à la fonction  $f$  implique immédiatement

$$\begin{aligned} NFA_b(|D|, \delta(v_D), \sigma) &= \max_m |\mathcal{D}_{\text{BR}}| C_{|\mathcal{D}_{\text{BR}}|}^{|D|} \cdot f(|D| - |\mathcal{L}|, \delta(v_D), \sigma, m) \\ &= |\mathcal{D}_{\text{BR}}| C_{|\mathcal{D}_{\text{BR}}|}^{|D|} \cdot f(|D| - |\mathcal{L}|, \delta(v_D), \sigma, 0) \\ &= NFA_a(|D|, \delta(v_D), \sigma). \end{aligned} \tag{7.3.36}$$

Les hypothèses  $H_0^a$  et  $H_0^b$  aboutissent au même nombre de fausses alarmes, noté  $NFA$ , d'où l'énoncé de la proposition 7.3.2.  $\square$

Ce résultat signifie simplement que le modèle naïf  $H_0(0)$  est le modèle le plus *contraire* de la famille de modèles  $H_0(m)$ , où  $m \in \mathbb{R}^{|\mathcal{D}_{\text{BR}}|}$ . En effet, en considérant le vecteur des moyennes  $m = 0$ , le rejet du modèle  $H_0(0)$  implique le rejet de tous les modèles naïfs de la famille  $H_0(m)$ .

En considérant le modèle naïf selon lequel l'image **BR** est un champ aléatoire de variables gaussiennes i.i.d.  $\mathcal{N}(0, \sigma^2)$  où  $\sigma > 0$  fixé, le nombre de fausses alarmes obtenu est alors indépendant du choix de la moyenne  $m$ .

Sous la forme donnée par l'équation (7.3.35), le nombre de fausses alarmes peut être calculé numériquement à l'aide d'un développement asymptotique. Cependant, pour une bonne performance de la méthode, il est nécessaire de pouvoir le calculer avec une meilleure précision. C'est la raison pour laquelle nous préférons utiliser l'expression suivante.

**Corollaire 7.3.4** *Sous les hypothèses  $H_0^a$  ou  $H_0^b$ ,*

$$NFA(|D|, \delta(v_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \Gamma_{inc} \left( \frac{|D| - |\mathcal{L}|}{2}, \frac{\delta(v_D)^2}{2\sigma^2} \right), \quad (7.3.37)$$

où, pour tout  $x \geq 0$  et  $a > 0$ ,  $\Gamma_{inc}(a, x) = \frac{1}{\Gamma(a)} \int_0^x e^{-t} t^{a-1} dt$ .

**Démonstration.**

D'après la proposition 7.3.2, le nombre de fausses alarmes obtenu pour les hypothèses  $H_0^a$  ou  $H_0^b$  est défini par

$$NFA(|D|, \delta(v_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \bar{f}(|D| - L, \delta(v_D), \sigma), \quad (7.3.38)$$

où pour tout  $q \in \mathbb{N}^+$ ,  $\delta \geq 0$ ,  $\sigma > 0$ ,

$$\begin{aligned} \bar{f}(q, \delta, \sigma) &\equiv f(q, \delta, \sigma, 0) \\ &= \frac{1}{\sigma^q (2\pi)^{\frac{q}{2}}} \int_{(x_1, \dots, x_q) \in \mathcal{B}_q(\delta)} e^{-\frac{1}{2} \sum_{i=1}^q (\frac{x_i}{\sigma})^2} dx_1 \cdots dx_q. \end{aligned}$$

Avec le changement de variables  $r^2 = \sum_{i=1}^q x_i^2$ , nous obtenons l'expression :

$$\bar{f}(q, \delta, \sigma) = \frac{C}{\sigma^q (2\pi)^{\frac{q}{2}}} \int_0^\delta e^{-\frac{r^2}{2\sigma^2}} r^{q-1} dr, \quad (7.3.39)$$

où  $C = q \times \text{Vol}(\mathcal{B}_q(1))$  et  $\text{Vol}(\mathcal{B}_q(1))$  représente le volume de la boule de rayon 1 en dimension  $q$ , i.e.

$$\text{Vol}(\mathcal{B}_q(1)) = \frac{\pi^{\frac{q}{2}}}{\Gamma(\frac{q}{2} + 1)},$$

où  $\Gamma$  est la fonction d'Euler, définie pour tout  $x \in \mathbb{R}^{+*}$  par

$$\Gamma(x) = \int_0^{+\infty} e^{-t} t^{x-1} dt.$$

Le changement de variable  $s = \frac{r^2}{2\sigma^2}$  permet d'obtenir

$$\bar{f}(q, \delta, \sigma) = \frac{C}{\sigma^q (2\pi)^{\frac{q}{2}}} \int_0^{\frac{\delta^2}{2\sigma^2}} e^{-s} (\sqrt{s})^{q-1} (\sigma\sqrt{2})^{q-1} \frac{\sigma\sqrt{2}}{2\sqrt{s}} ds \quad (7.3.40)$$

$$= \frac{C}{2\pi^{\frac{q}{2}}} \int_0^{\frac{\delta^2}{2\sigma^2}} e^{-s} s^{\frac{q}{2}-1} ds \quad (7.3.41)$$

$$= \frac{q\pi^{\frac{q}{2}}}{2(\pi)^{\frac{q}{2}}\Gamma(\frac{q}{2} + 1)} \int_0^{\frac{\delta^2}{2\sigma^2}} e^{-s} s^{\frac{q}{2}-1} ds \quad (7.3.42)$$

$$= \frac{q}{2\Gamma(\frac{q}{2} + 1)} \int_0^{\frac{\delta^2}{2\sigma^2}} e^{-s} s^{\frac{q}{2}-1} ds. \quad (7.3.43)$$

La fonction  $\Gamma$  vérifie, pour tout  $q \in \mathbb{R}^{+*}$ ,

$$\Gamma\left(\frac{q}{2} + 1\right) = \frac{q}{2} \Gamma\left(\frac{q}{2}\right).$$

Donc

$$\bar{f}(q, \delta, \sigma) = \frac{1}{\Gamma(\frac{q}{2})} \int_0^{\frac{\delta^2}{2\sigma^2}} e^{-s} s^{\frac{q}{2}-1} ds, \quad (7.3.44)$$

où la fonction  $\Gamma$  d'Euler vérifie, pour tout  $q \in \mathbb{N}$ ,

$$\Gamma(\frac{q}{2}) = \begin{cases} (\frac{q}{2} - 1)! & \text{si } q > 1 \text{ est pair} \\ \frac{(q-2)(q-4) \cdots 6 \cdot 4 \cdot 2}{2^{\frac{q-1}{2}}} \sqrt{\pi} & \text{si } q > 2 \text{ est impair} \\ \sqrt{\pi} & \text{si } q = 1. \end{cases} \quad (7.3.45)$$

La fonction  $\bar{f}$  s'exprime à partir de la fonction *gamma incomplète*, notée  $\Gamma_{inc}$  et définie pour tout  $a > 0$  et  $x \geq 0$  par

$$\Gamma_{inc}(a, x) = \frac{1}{\Gamma(a)} \int_0^x e^{-t} t^{a-1} dt. \quad (7.3.46)$$

Le nombre de fausses alarmes est alors défini par

$$NFA(|D|, \delta(v_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \Gamma_{inc} \left( \frac{|D| - |\mathcal{L}|}{2}, \frac{\delta(v_D)^2}{2\sigma^2} \right).$$

□

Le critère obtenu dépend donc uniquement de la taille du sous-domaine considéré, de l'erreur calculée sur ce sous-domaine et de l'écart-type du modèle naïf  $H_0$ . Les deux premiers paramètres mentionnés sont directement déterminés à partir des images étudiées. Le seul paramètre restant est alors  $\sigma$ , l'écart-type du modèle naïf. Ce dernier paramètre est finalement le seul paramètre arbitraire du modèle. Il peut être fixé comme un paramètre physique *a priori*, envisagé par exemple comme une mesure de séparabilité ou plus généralement comme un moyen d'introduire de l'information *a priori*. Dans le cadre de notre étude, nous proposons de fixer la variance  $\sigma^2$  du modèle naïf égale à la variance empirique de l'image observée. Ce choix permet de garantir, en moyenne, l'absence de toute détection dans une image de bruit blanc et de respecter ainsi le principe de Helmholtz. En effet, une image de bruit blanc correspond précisément au modèle  $H_0$  pour l'écart-type empirique de l'image. Ce résultat est abordé de manière analytique avec la proposition 9.2.1. De plus, ce dernier paramètre est alors à son tour directement déterminé à partir des images considérées, ce qui rend, en pratique, la méthode entièrement automatique et, grâce à la formulation (7.3.37), le nombre de fausses alarmes numériquement calculable.

Avec le modèle  $H_0$ , l'hypothèse d'indépendance des pixels revient à considérer que la probabilité d'occurrence d'une image **BR** est égale au produit des probabilités d'occurrences en chaque pixel. Par conséquent, ce modèle revient à envisager toutes les permutations de  $D$  de manière équiprobable. De plus, fixer le vecteur des moyennes  $m$  revient à se donner la loi de l'histogramme de l'image. Ce choix est arbitraire. La possibilité de prendre pour modèle *a contrario* l'hypothèse suivante a aussi été envisagée.

**Définition 7.3.5 (modèle *a contrario*  $H'_0$ )** Le modèle naïf ( $H'_0$ ) pour l'image **BR** est un champ aléatoire de variables i.i.d. suivant la loi donnée par l'histogramme mesuré empiriquement sur l'image.

Sous cette hypothèse, c'est davantage l'indépendance des pixels basse résolution qui est testée. Le principe de la méthode est alors d'estimer un histogramme d'après les proportions et les moyennes des classes, puis de définir un sous-domaine comme significatif si l'histogramme estimé sur ce sous-domaine est *étonnamment* en adéquation avec le modèle *a contrario* (i.e. l'histogramme mesuré empiriquement sur l'image). Plus précisément, si  $A$  représente la matrice des proportions  $(\alpha_l(y))_{y,l}$ , il s'agit de considérer la probabilité

$$\mathbb{P}_{H'_0}[\delta_\mu(V_D) \geq \delta_\mu(v_D)] = \mathbb{P}_{H'_0}[v_D \in \mathcal{B}(A\mu, \delta_\mu(V_D))] \quad (7.3.47)$$

$$= \int_{\mathcal{B}(A\mu, \delta_\mu(V_D)) \times \mathbb{R}^{p-q}} \prod_{i=1}^{|D|} dh(x_i), \quad (7.3.48)$$

où  $dh$  représente la mesure associée à l'histogramme empirique de l'image. Cependant, dans notre contexte, les moyennes  $\mu$  des classes sont inconnues. C'est la raison pour laquelle, comme précédemment (cf. (7.2.28)), nous nous intéressons plutôt à l'erreur **minimale**  $\delta = \min_\mu \delta_\mu$  et, donc, à la probabilité

$$\mathbb{P}_{H'_0}[\delta(V_D) \geq \delta(v_D)] = \mathbb{P}_{H'_0}[\|PV_D\| \geq \delta(v_D)]. \quad (7.3.49)$$

Déterminer cette loi, c'est être capable de compter le nombre de points de la forme  $\lambda_{i_1}, \dots, \lambda_{i_p}$  dans le cylindre  $\mathcal{B}(A\mu, \delta_\mu(V_D)) \times \mathbb{R}^{p-q}$ . Remarquons que si la loi donnée par l'histogramme de l'image est proche d'une loi gaussienne, le modèle  $H'_0$  est équivalent au modèle  $H_0$  pour le calcul du nombre de fausses alarmes. A défaut de calculer cette probabilité, une estimation de l'intégrale peut être donnée, par exemple, par la méthode de Monte Carlo [Liu, 2001, Rubinstein, 1981] et conduire à une estimation du *NFA* pour le modèle  $H'_0$ . La méthode de détection, entièrement basée sur l'estimation de la probabilité (7.3.49), serait alors difficile à analyser théoriquement, notamment pour en évaluer les seuils de détection. De plus, nous verrons dans la section 8.2, que l'algorithme de minimisation du nombre de fausses alarmes doit être itéré un grand nombre de fois. Il est donc important de pouvoir optimiser cette minimisation. Dans la suite de l'étude, nous considérons uniquement le modèle *a contrario* défini par l'hypothèse  $H_0$ , pour lequel le nombre de fausses alarmes a déjà été déterminé 7.3.4.

Nous avons exposé le modèle de détection dans le cas de la comparaison d'une image unique à la labellisation de référence. La généralisation au cas multidimensionnel est immédiate en considérant un vecteur d'images. Cependant, l'absence éventuelle de données en certains pixels de l'image ou la présence de données aberrantes sur différents pixels selon les dates d'acquisition nécessite de définir les domaines images correspondants à chaque date avant de définir, section 7.4, le *nombre de fausses alarmes* pour des séquences d'images.

## 7.4. Modèle multidimensionnel

Dans cette section, nous présentons la version multidimensionnelle du modèle *a contrario*, adaptée du cas monodimensionnel précédemment décrit pour l'application à des séquences d'images. Théoriquement, l'adaptation est directe. L'objectif est simplement de comparer une série d'images **BR**, multispectrale ou multitemporelle, à une labellisation de manière à détecter le sous-domaine sur lequel la labellisation est cohérente avec la séquence d'images. Le complémentaire de ce sous-domaine correspondrait alors à l'ensemble des pixels de changements apparus depuis la date de référence.

En pratique, l'utilisation d'images multispectrales est très simple dans la mesure où il s'agit d'acquisitions simultanées par un même satellite. Par définition, les images sont donc parfaitement superposables et définies sur un même domaine. Le cas des séquences multitemporelles est généralement plus délicat dès l'instant où le domaine de définition des images, *i.e.* l'ensemble des pixels dont la valeur est utilisable, peut varier selon les dates. Pour la détection de changements de types d'occupation du sol en particulier, notre objectif est centré sur la comparaison d'une séquence multitemporelle à une labellisation de référence. C'est donc le cas multitemporel que nous présentons ici, même si tout type de séquence pourrait être utilisé.

Supposons que toutes les images de la séquence sont parfaitement superposables, même si les valeurs peuvent être aberrantes ou manquantes sur certains pixels, selon les dates. En effet, lorsque des nuages apparaissent sur une partie de la scène, ils masquent, dans le cas des données optiques, le rayonnement émis par la surface. La valeur acquise par le capteur n'est alors pas utilisable. Dans le cas des données **SPOT/VGT**, les synthèses décennales sont réalisées à partir d'images acquises sur une période de dix jours, *i.e.* avec des conditions d'acquisition variables d'un pixel à l'autre en terme d'angle de visée et d'illumination. Les caractéristiques physiques des classes dépendent à la fois de l'angle d'observation et de l'angle solaire. Dans cette étude, nous faisons une hypothèse forte de stationnarité sur les caractéristiques des classes. En pratique, nous décidons d'éliminer les pixels pour lesquels les angles d'observation sont trop différents.

Soit  $\mathcal{T}$  l'ensemble des dates d'acquisition de chaque image de la séquence. Compte tenu de la présence éventuelle de données manquantes, le domaine spatio-temporel de notre étude est un sous-domaine de  $\mathcal{D}_{BR} \times \mathcal{T}$ , noté  $\Omega = \{(y, t) \text{ valides} \in \mathcal{D}_{BR} \times \mathcal{T}\}$ . Pour une date  $t$  fixée, l'ensemble des pixels valides forme alors un domaine spatial noté  $\Omega_t = \{y, (y, t) \in \Omega\}$ , et pour un pixel basse résolution  $y$  fixé, l'ensemble des dates valides forme un domaine temporel noté  $\Omega_y = \{t, (y, t) \in \Omega\}$ . La séquence d'images est alors notée  $v = (v_t)_{t \in \mathcal{T}}$ , où chaque image

$$\begin{aligned} v_t : \Omega_t &\rightarrow \mathbb{R} \\ y &\mapsto v_t(y). \end{aligned}$$

Notre objectif est de détecter dans quel sous-domaine spatio-temporel  $\omega \subset \Omega$  une labellisation donnée reste cohérente avec une séquence d'images. De même que dans le cas monodimensionnel, la série multidimensionnelle est modélisée en fonction de la composition des pixels en termes de

classes et des valeurs caractéristiques (moyennes) de chaque classe. Les taux d'occupation de chaque classe  $l$  dans un pixel  $y$ , notés  $\alpha_l(y)$  et calculés à partir de la labellisation **HR**, font intervenir uniquement les propriétés géométriques des images. En supposant que les images de la séquence sont parfaitement recalées et en l'absence de changements, les taux d'occupation ne varient pas au cours du temps. Ils peuvent donc être calculés à partir de la labellisation et d'une image quelconque de la série. En revanche, le signal émis par les différents types de couvert évolue et le vecteur  $\mu$  des moyennes par label dépend alors de la date considérée ; soit  $\mu = (\mu_t(l))_{l \in \mathcal{L}, t \in \mathcal{T}}$ . Etant données les moyennes  $\mu$ , l'intensité observée à la date  $t$  en un pixel  $y \in \Omega_t$  peut être estimée selon le modèle linéaire de mélange par

$$\hat{v}_t(y, \mu) = \sum_{l \in \mathcal{L}} \alpha_l(y) \mu_t(l). \quad (7.4.50)$$

Soient  $\omega$  un sous-domaine du domaine spatio-temporel des pixels valides  $\Omega$ ,  $v_\omega = \{v_t(y), (y, t) \in \Omega\}$  la séquence d'images restreinte au domaine spatio-temporel valide et  $V_\omega = \{V_t(y), (y, t) \in \Omega\}$  le champ aléatoire dont elle est une réalisation. La différence entre l'intensité observée sur un sous-domaine spatio-temporel et son estimation peut alors être mesurée, comme en 7.1.2, par la norme euclidienne

$$\delta_\mu(v_\omega) = \sqrt{\sum_{(y,t) \in \omega} (v_t(y) - \hat{v}_t(y, \mu))^2}, \quad (7.4.51)$$

et l'erreur minimale sur le domaine spatio-temporel

$$\delta(v_\omega) = \min_{\mu \in \mathbb{R}^{|\mathcal{L}| \times |\mathcal{T}|}} \delta_\mu(v_\omega). \quad (7.4.52)$$

Le principe de la méthode est alors identique au cas monodimensionnel, en considérant le sous-domaine spatio-temporel  $\omega$  au lieu du sous-domaine spatial  $D$  et le modèle *a contrario* suivant.

**Définition 7.4.1 (modèle a contrario)** *Le modèle a contrario pour la séquence d'images **BR** est un champ aléatoire  $V$  de  $|\Omega|$  variables gaussiennes i.i.d.  $\mathcal{N}(0, \sigma^2)$  où  $\sigma > 0$  fixé.*

En pratique, le paramètre  $\sigma^2$  est fixé, égal à la variance empirique de la séquence observée. Remarquons que le temps et l'espace jouent un rôle équivalent ici, puisque nous ne considérons que des variables aléatoires indépendantes. Le nombre de fausses alarmes défini dans le cas de monodimensionnel dépend uniquement de trois paramètres : le cardinal du sous-domaine considéré, le résidu associé et la variance empirique de l'image. De manière similaire, il est déterminé dans le cas de séquences d'images par le théorème suivant.

**Proposition 7.4.2** *Avec le modèle a contrario donné par la définition 7.4.1, le nombre de fausses alarmes associé à un sous-domaine spatio-temporel  $\omega$  de  $\Omega$  d'une séquence  $v$  de  $T$  images est déterminé par*

$$NFA(\omega, \delta(v_\omega), \sigma) = |\Omega| C_{|\Omega|}^{|\omega|} \cdot \Gamma_{inc} \left( \frac{|\omega| - L \times T}{2}, \frac{\delta^2(v_\omega)}{2\sigma^2} \right), \quad (7.4.53)$$



où, pour tout  $x \geq 0$  et  $a > 0$ ,

$$\Gamma_{inc}(a, x) = \frac{1}{\Gamma(a)} \int_0^x e^{-t} t^{a-1} dt. \quad (7.4.54)$$

#### Démonstration.

Evident d'après le théorème 7.2.5. □

La recherche du sous-domaine qui minimise le nombre de fausses alarmes permet alors d'accéder directement au sous-domaine des changements. Cependant, si des informations peuvent être inutilisables sur une zone de l'image à une date donnée, cette zone peut varier à chaque date. Les domaines de définition de chaque image de la séquence sont alors différents, et l'algorithme de détection présente quelques différences avec celui du cas monodimensionnel (cf. Chapitre 8).

Bien qu'il s'agisse de séquences temporelles d'images, les propriétés spatiales et temporelles des données ne sont pas prises en compte dans ce modèle. Il ne nous semble pas raisonnable d'introduire de terme de régularité spatiale dans la mesure où notre objectif est conditionné par l'utilisation de données **BR** où chaque pixel représente une surface d'environ  $1\text{km}^2$  au sol. Ce modèle permet, indifféremment, l'utilisation de séries temporelles, d'images multispectrales ou de séries temporelles multispectrales. Il revient à tester la valeur d'un pixel par rapport aux autres pixels de l'image à une date donnée plutôt que par rapport aux pixels d'une date précédente. Un changement tel que l'inversion des occupations du sol de deux parcelles de même taille, inférieure à la surface représentée dans un pixel **BR**, ne devrait donc pas être détectable par cette méthode. Finalement, le test porte sur la cohérence interne d'une image et non relativement aux autres dates de la séquence, ce qui devrait rendre la mesure de cohérence peu sensible aux variations d'intensités entre les images de la séquence étudiée. Les performances du modèle sont analysées théoriquement dans le chapitre 9.

La section suivante présente quelques éléments de comparaison entre le calcul du nombre de fausses alarmes et les procédures statistiques classiques.

## 7.5. Comparaison avec des procédures statistiques classiques

Dans cette section, nous tentons de situer le calcul du *nombre de fausses alarmes* dans un cadre statistique classique. Dans un premier temps, nous présentons certains aspects du problème des tests multiples en relation avec l'approche du *NFA*. Ensuite, nous proposons d'envisager le problème de la détection de changements sous la forme d'une analyse de la variance.

### 7.5.1. Les procédures de tests multiples

Le calcul du nombre de fausses alarmes est lié au problème statistique des tests multiples [Hochberg et Tamhane, 1987] dans la mesure où il tient compte du nombre de tests réalisés dans sa définition. Les procédures de tests multiples ont été développées dans l'objectif de traiter des

grands nombres de données sans augmenter le nombre moyen d'erreurs de détection. La démarche statistique classique pour tester  $N$  variables aléatoires tout en contrôlant le taux de faux positifs serait alors de seuiller les  $N$  p-valeurs individuellement à un niveau  $\alpha$ , garantissant ainsi un nombre moyen de faux positifs inférieur à  $\alpha N$ , *i.e.* croissant en fonction du nombre de tests. En supposant que les  $N$  tests sont indépendants, la probabilité de faire au moins une erreur de première espèce vaut alors  $1 - (1 - \alpha)^N$ , soit par exemple 40% lorsqu'il s'agit de faire 10 tests au niveau 5%.

En ce qui concerne le problème de détection de changements que nous étudions, il pourrait être envisagé sous la forme d'un test d'adéquation en se donnant un modèle théorique pour l'image **BR**. Par exemple, le modèle défini dans la Partie II où les variables aléatoires associées à chaque pixel sont supposées gaussiennes, paramétrées selon le pourcentage d'occupation de chaque classe dans le pixel et les caractéristiques de ces classes, peut être considéré comme modèle *a priori*. Les tests des  $2^{|\mathcal{D}_{BR}|}$  sous-domaines du domaine  $\mathcal{D}_{BR}$  sont alors considérés selon les hypothèses :

$H_0(D)$  : "la restriction de l'image **BR** au sous-domaine  $D$  n'est pas structurée selon la classification de référence."

$H_1(D)$  : "la restriction de l'image **BR** au sous-domaine  $D$  est structurée selon la classification de référence."

Les pixels détectés comme représentants des changements sont alors ceux qui satisfont l'hypothèse nulle selon laquelle l'image observée suit la loi théorique. Ce test peut être réalisé, par exemple, avec une statistique du  $\chi^2$  [Saporta, 1990]. Avec les notations des sections précédentes, la règle de décision consiste alors à rejeter l'hypothèse  $H_0(D)$  si  $\delta(V_D) \geq \delta(v_D)$ . Si chaque test de  $H_0(D)$  par  $H_1(D)$  est de niveau  $\alpha_D$ , la probabilité de rejeter  $H_0(D)$  à tort satisfait

$$\mathbb{P}(\delta(V_D) \geq \delta(v_D) | H_0(D)) \leq \alpha_D. \quad (7.5.55)$$

Après avoir réalisé  $N$  tests d'hypothèse  $H_0 = \cup_{D \in \mathcal{D}_{BR}} H_0(D)$  contre  $H_1 = H_0^c$ , il semble naturel de rejeter  $H_0$  pour  $H_1$  si il existe un sous-domaine  $D \in \mathcal{D}_{BR}$  tel que l'hypothèse  $H_0(D)$  soit rejetée pour  $H_1(D)$ , *i.e.* tel que  $\delta(V_D) \geq \delta(v_D)$ . Remarquons que la probabilité de rejeter  $H_0$  à tort satisfait alors

$$\begin{aligned} \mathbb{P}(\exists D \in \mathcal{D}_{BR} ; \delta(V_D) \geq \delta(v_D) | H_0) &\leq \sum_{D \in \mathcal{D}_{BR}} \mathbb{P}(\delta(V_D) \geq \delta(v_D) | H_0^D) \\ &\leq \sum_{D \in \mathcal{D}_{BR}} \alpha_D. \end{aligned} \quad (7.5.56)$$

L'approche naïve évoquée précédemment serait de considérer que chaque test est réalisé au niveau  $\alpha_D = \alpha$ , indépendamment du domaine testé et, par conséquent, que la probabilité de rejeter  $H_0$  à tort est inférieure à  $N\alpha$  (soit  $2^{10\,000} \times 0.05$  pour une image de taille  $100 \times 100$  et un niveau de significativité  $\alpha = 0.05$ ). Pour éviter une telle augmentation des erreurs de première espèce, la procédure de [Bonferroni, 1936] a pour objectif de maintenir la probabilité d'obtenir une fausse alarme à un niveau inférieur à  $\alpha > 0$  en testant, par exemple, chaque sous-domaine à un niveau  $\alpha_D \leq \frac{\alpha}{N}$ . L'hypothèse  $H_0$  est alors rejetée si  $\alpha_D \leq \frac{\alpha}{N}$  pour tout  $D \in \mathcal{D}_{BR}$ . Cette stratégie est très contraignante puisque chaque test individuel est alors maintenu à un seuil particulièrement

exigeant de manière à contrôler le taux de fausses alarmes, au détriment des erreurs de seconde espèce alors augmentées. Cette procédure a donc tendance à manquer les détections. Formellement, le calcul du nombre de fausses alarmes (défini pour  $N = 2^{|\mathcal{D}_{BR}|}$  tests) est équivalent à la procédure de Bonferroni puisqu'il revient simplement à tester si  $N\alpha_D \leq \alpha$ . Cependant, leur concept est différent et, en particulier, le *NFA* permet directement de contrôler le nombre de fausses alarmes (faux positifs). En effet, contrairement à la procédure de Bonferroni, il a un sens intrinsèque permettant, par exemple, de décider d'accepter en moyenne exactement 10 erreurs sur un million de tests (alors qu'un test de niveau 10 n'a pas de sens). La définition du *NFA* que nous utilisons en pratique en est une variante puisque le nombre de tests considéré ( $N = |\mathcal{D}_{BR}|C_{|\mathcal{D}_{BR}|}^{|D|}$ ) pour le *NFA* pénalise les sous-domaines testés selon leur taille, en favorisant ceux de grande et de petite taille. Formellement, il s'agit de tester tous les sous-domaines  $D$  au niveau  $\alpha_D$  tel que  $\alpha_D \leq \frac{\alpha}{|\mathcal{D}_{BR}|C_{|\mathcal{D}_{BR}|}^{|D|}}$ , choix qui permet de garantir une erreur de première espèce inférieure à  $\alpha$  puisque

$$\sum_{D \in \mathcal{D}_{BR}} \alpha_D \leq \sum_{D \in \mathcal{D}_{BR}} \frac{\alpha}{|\mathcal{D}_{BR}|C_{|\mathcal{D}_{BR}|}^{|D|}} \leq \alpha. \quad (7.5.57)$$

De nombreuses procédures ont été proposées dans la littérature en tant qu'alternatives à celle de Bonferroni [Hommel, 1988, Hochberg, , Benjamini et Hochberg, 1995], moins conservatrices. Elles sont basées sur le contrôle de différents types d'erreurs telles que, par exemple, le FWER (*Family Wise Error Rate*) qui permet de contrôler la probabilité d'avoir au moins un faux positif, le PCER (*Per Comparison Error Rate*) qui contrôle l'espérance de la proportion de faux positifs parmi le nombre total de tests, ou encore le FDR (*False Discovery Rate*) pour contrôler la proportion attendue d'hypothèses rejetées à tort. En particulier, la procédure de [Benjamini et Hochberg, 1995], très utilisée, permet un contrôle du FDR tout en considérant les tests dans leur globalité. La décision est alors prise en considérant un ensemble de tests significatifs et non pas chaque test individuellement. Plus précisément, cette procédure consiste, dans un premier temps, à ordonner les p-valeurs

$$p_{(1)} \leq \dots \leq p_{(i)} \leq \dots \leq p_{(N)}$$

puis, si  $k = \max\{i : p_{(i)} \leq \frac{i}{N}\alpha\}$  existe, à rejeter  $H_0(i)$  pour  $i = 1 \dots k$  et garantir ainsi un taux de fausses découvertes  $FDR \leq \alpha$ . Cette procédure est moins conservatrice que celle de Bonferroni puisqu'elle permet d'accepter tous les sous-domaines pour lesquels les p-valeurs sont inférieures à  $p_k$  alors que la correction de Bonferroni ne permet d'accepter que les sous-domaines dont la p-valeur est inférieure au seuil  $\frac{\alpha}{N}$ . Finalement, la procédure de [Benjamini et Hochberg, 1995] permet de détecter globalement un ensemble de sous-domaines ne représentant pas de changements, tout en acceptant un certain pourcentage d'erreurs, mais aucun sous-domaine n'est détecté individuellement. Dans le cadre de notre étude, il semble difficile de considérer l'union de sous-domaines significatifs et la détection d'une collection de sous-domaines comme étant significatifs est difficile d'interprétation. La procédure de [Benjamini et Hochberg, 1995] ne paraît donc pas adaptée au problème que nous considérons. De plus, les valeurs de *NFA* correspondent typiquement à des probabilités très faibles (de l'ordre de  $10^{-1000}$ ). Un seuil du type  $\frac{\alpha}{N}$  sur le *NFA* ne serait alors pas contraignant.

La pertinence de la définition du  $NFA$  repose sur le résultat selon lequel l'espérance du nombre de fausses alarmes d'événements  $\varepsilon$ -significatifs est inférieur à  $\varepsilon$ . Cette propriété permet de fixer automatiquement les seuils de décision de manière à garantir un nombre moyen de fausses alarmes fixé *a priori* (par exemple,  $\varepsilon = 1$ ). De plus, nous décidons de retenir le domaine le plus significatif, *i.e.* celui qui minimise le  $NFA$ , comme domaine le plus cohérent avec la classification.

### 7.5.2. L'analyse de variance

Le modèle *a contrario* ( $H_0(m)$ ) que nous considérons consiste, en un sens, à supposer que l'image n'est pas structurée (elle correspond à une classe unique) et à s'étonner de lui découvrir une structure significative, correspondant à la classification de référence. Finalement, cette approche repose sur l'étude des variances *intra-classes* (les erreurs quadratiques) relativement à la variance de l'image. Ce principe rappelle celui de l'analyse de variance (ANOVA) [Saporta, 1990] couramment utilisée en statistique. Dans cette section, nous proposons d'envisager la détection de changements sous la forme d'une analyse de variance dans le but de comparer les deux approches (ANOVA et  $NFA$ ).

L'analyse de variance (ANOVA) est une famille de modèles statistiques développée par R.A. Fisher dans les années 1920-1930 dont l'objectif est d'étudier l'influence d'un ou de plusieurs facteurs sur une variable quantitative. Cette méthode est très utilisée dans les domaines biologiques, médicaux et pharmaceutiques pour analyser la significativité des effets d'un ou de plusieurs facteurs (par exemple, des traitements) sur autant d'échantillons. Il existe différents types de modèles selon le nombre de facteurs étudiés, et la nature des modalités associées au facteur (on parle de modèle à effets fixes, aléatoires ou mixtes). Dans tous les cas, l'analyse de variance peut être vue comme une comparaison multiple de moyennes de différents échantillons constitués par les différentes modalités des facteurs. Pour notre étude, nous nous intéressons plus particulièrement au cas où un seul paramètre (appelé facteur) est susceptible d'influer sur les données étudiées. Supposons que ce facteur puisse admettre  $k$  valeurs différentes (les modalités). L'échantillon de données global, de taille  $n$ , peut être décomposé en  $k$  échantillons de taille  $n_1, n_2, \dots, n_k$ . Dans le cas du modèle à effets fixes, les variables aléatoires qui modélisent les données correspondant à la modalité  $k$  sont supposées indépendantes et de même loi  $\mathcal{N}(\mu_k, \sigma^2)$ . L'objectif est alors de déterminer si la variabilité observée dans les données est uniquement due au hasard ou s'il existe effectivement des différences significatives entre les classes, qui soient imputables au facteur. Il s'agit donc de comparer les variances empiriques de chaque échantillon à la variance de l'échantillon global.

Le principe de la méthode de détection de changements que nous proposons est, en un sens, assez proche d'une analyse de la variance à effets fixes, à un facteur. Cependant, ce lien ne peut être établi que dans le contexte où la classification et l'image observée (ou la séquence) ont la même résolution (*i.e.* sans problème de désaggrégation de pixels mixtes). En effet, le problème de la détection de changements **HR** peut être abordé en cherchant à savoir si la classification de référence a un *effet* sur l'image observée. Si la classification contient  $L$  labels, l'image est constituée d'un ensemble de  $L$  échantillons de tailles  $n_1, \dots, n_L$  correspondant chacun à un label

donné. Chaque pixel de l'image est associé à un seul label, donc  $n_1 + \dots + n_L = n$ , où  $n$  représente le nombre total de pixels (l'échantillon global). L'analyse de la variance (ANOVA à 1 facteur) permet d'évaluer les effets des  $L$  labels, comme s'ils correspondaient à des traitements, sur les valeurs observées des  $n_1, \dots, n_L$  pixels. Plus précisément, l'analyse de la variance consiste à tester si la variabilité observée dans les données est uniquement due au hasard ou s'il existe effectivement des différences significatives entre les échantillons, explicables par la classification. Les variables aléatoires modélisant les données correspondantes à chaque label sont supposées indépendantes et de même loi  $\mathcal{N}(\mu_l, \sigma^2)$  pour le label  $l$ . Il s'agit de comparer les variances empiriques de chaque échantillon à la variance globale de l'image. Les variances empiriques sont de deux types : la variance *intra-classes* ( $\sigma_{intra}^2$ ) qui décrit la variabilité à l'intérieur d'une même classe et la variance *inter-classes* ( $\sigma_{inter}^2$ ) qui décrit les différences entre les classes et peut éventuellement être expliquée par la classification. La variance globale de l'image vaut  $\sigma_{intra}^2 + \sigma_{inter}^2$ . Le test ANOVA consiste à accepter l'effet de la classification sur l'image observée si le rapport pondéré de la variance résiduelle (*inter-classes*) sur la variance expliquée (*intra-classes*) est significativement trop grand. Sous l'hypothèse

$$H_0 : \mu_1 = \dots = \mu_L,$$

la variable

$$R = \frac{\sigma_{intra}^2 / (L - 1)}{\sigma_{inter}^2 / (n - L)}$$

suit la loi de Fisher  $\mathcal{F}(L - 1, n - L)$  et le test ANOVA permet de rejeter l'hypothèse nulle d'égalité des moyennes (*i.e.* d'accepter les effets de la classification) quand le rapport  $R$  est trop grand par rapport aux quantiles de la loi  $\mathcal{F}(L - 1, n - L)$ .

Le modèle *NFA* décrit section 7.2 est défini à partir de l'hypothèse  $H_0$  et la mesure  $\delta$ . En supposant que la classification et l'image observée sont de même résolution (*i.e.* tous les pixels sont purs), l'erreur quadratique  $\delta^2(v_D)$  mesurée entre l'observation et son estimation peut être vue comme la variance *inter-classes* du modèle ANOVA (variance résiduelle). Il revient donc à comparer la variance *intra-classes* de chaque échantillon à la variance de l'échantillon global ( $\delta^2(v_D)/\sigma^2$ ) alors que le principe de l'analyse de variance repose sur la comparaison des variances empiriques *intra-classes* et *inter-classes* à la variance de l'échantillon global. Dans le cadre de la détection de changements à partir d'images **BR**, cette approche n'est pas directement utilisable. En effet, l'analyse de la variance suppose de pouvoir décomposer un échantillon global en différents groupes définis selon les modalités du facteur considéré (ici, selon les classes). Cette décomposition n'est pas possible à partir de données **BR** dans la mesure où la valeur en chaque pixel résulte alors des contributions de plusieurs modalités.

La modélisation *a contrario* a prouvé son efficacité dans le cadre de diverses applications de détection (*cf.* [Desolneux et al., 2003, Dibos et al., 2005, Cao et al., 2005]). Son utilisation est motivée en grande partie par le fait qu'elle conduise à des méthodes de détection réduites à un unique paramètre : le nombre moyen de fausses détection autorisé. L'objectif de cette partie est d'explorer son potentiel pour la détection de changements de taille sous-pixellique à partir de séquences d'images. Pour cela, nous commençons par aborder certains aspects numériques liés à la

mise en oeuvre du modèle *a contrario* défini au cours de ce chapitre (*cf.* chapitre 8). En particulier, nous décrivons les algorithmes monotemporel et multitemporel adoptés pour la recherche du sous-domaine spatial sur lequel la séquence d'images est la plus cohérente avec la labellisation.



## 8. Aspects numériques

La méthode de détection de changements que nous proposons est basée sur la définition d'un critère appelé *nombre de fausses alarmes* pour lequel nous construisons un algorithme suivant le principe de l'échantillonnage aléatoire (**RANSAC**) introduit en analyse d'images par [Fischler et Bolles, 1981].

### 8.1. Problématique

Etant données une labellisation **HR** et une image ou une séquence **BR** de la même scène, le *nombre de fausses alarmes* est défini, dans le cas monodimensionnel, pour tout sous-domaine spatial du domaine image par

$$NFA(|D|, \delta(v_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \Gamma_{inc}\left(\frac{|D| - |\mathcal{L}|}{2}, \frac{\delta(v_D)^2}{2\sigma^2}\right),$$

où, pour tout  $x \geq 0$  et  $a > 0$ ,  $\Gamma_{inc}(a, x) = \frac{1}{\Gamma(a)} \int_0^x e^{-t} t^{a-1} dt$ . Dans le cas multitemporel, une expression analogue est obtenue pour tout sous-domaine spatio-temporel (cf. proposition 7.4.2). Un sous-domaine est alors détecté comme cohérent avec une labellisation donnée s'il correspond au sous-domaine qui minimise le *NFA*. Ce dernier dépend de la taille de l'image **BR**, du nombre de labels, de la taille du sous-domaine  $D$  considéré, de l'écart-type du modèle naïf et du résidu quadratique cumulé sur ce sous-domaine

$$\delta^2(v_D) = \sum_{y \in D} \left( v(y) - \sum_{l \in \mathcal{L}} \alpha_l(y) \mu(l) \right)^2.$$

Le choix de l'écart-type du modèle naïf est discuté dans la section 7.3, il est fixé égal à la variance empirique de l'image **BR**. De cette manière, nous nous attendons, en particulier, à ce que la méthode ne détecte rien dans une image de bruit blanc. Tous les paramètres du *NFA* s'obtiennent donc directement à partir des données, à l'exception du résidu quadratique cumulé qui dépend de la famille  $\{\mu(l)\}_{l \in \mathcal{L}}$  des moyennes des classes, inconnue *a priori*. Il est alors nécessaire d'estimer les moyennes  $\mu(l)$  caractéristiques de chaque label  $l$  avant de pouvoir calculer les résidus quadratiques pour un sous-domaine considéré, puis le *NFA* qui lui est associé. L'estimation de cette famille de paramètres et de la détection sont deux problèmes étroitement liés dans le sens où la qualité de l'estimation influence nettement celle de la détection.

Les méthodes classiques d'estimation telles que la méthode des moindres carrés cherchent à optimiser, pour une distance donnée, l'adéquation entre un modèle défini et les données. Par



exemple, la méthode des moindres carrés estimerait la famille  $\mu$  en résolvant le problème

$$\min_{\mu \in \mathbb{R}^{\mathcal{L}}} \sum_{y \in D} \left( v(y) - \sum_{l \in \mathcal{L}} \alpha_l(y) \mu(l) \right)^2.$$

Ce type de méthodes est extrêmement sensible à la présence de points aberrants dans la mesure où leur objectif est de s'approcher autant que possible de la totalité de l'échantillon, partant du principe qu'il y aura toujours assez de bonnes valeurs dans l'échantillon pour moyenniser l'effet des éventuelles grandes déviations. La présence d'un simple point aberrant qui serait loin de la solution suffirait à biaiser fortement le résultat final. Dans le contexte de la détection de changements en particulier, les pixels correspondants à des changements jouent le rôle de points aberrants dans le sens où leur utilisation pour l'estimation des moyennes risque de la perturber fortement. Les méthodes robustes, telles que les M-estimateurs, LMedS ou RANSAC ont été introduites dans l'objectif de faire face à ce problème. La méthode des M-estimateurs [Zhang et al., 1994], par exemple, permet une bonne estimation, même en présence de 50% de pixels incorrects, en pondérant la distance considérée. Le choix de la pondération reste, cependant, délicat à déterminer. Le principe de la méthode RANSAC (*Random Sampling Consensus*), proposée par [Fischler et Bolles, 1981], est d'utiliser un sous-ensemble de l'échantillon aussi petit que possible et de le compléter avec des données consistantes lorsque c'est possible plutôt que d'utiliser le plus de données possibles pour obtenir une solution et d'en éliminer ensuite les mauvais points. Plus précisément, supposons que l'on sélectionne au hasard un sous-ensemble des données et que nous estimions les paramètres à partir de ce sous-ensemble. Par chance, il peut arriver que le sous-ensemble considéré ne contienne aucun point incorrect, l'estimation des paramètres est alors correcte et permet de partitionner l'ensemble des données entre points corrects et incorrects. La stratégie RANSAC est basée sur l'idée selon laquelle répéter cette procédure d'échantillonnage aléatoire un grand nombre de fois doit permettre d'obtenir une solution satisfaisante. Cette approche, introduite en analyse d'images par [Fischler et Bolles, 1981], permet une bonne estimation même en présence de nombreux points incorrects (environ 50%). Cependant, elle est limitée, à son tour, par le choix arbitraire d'un seuil à partir duquel décider qu'un sous-ensemble est en adéquation avec le modèle considéré.

Par ailleurs, la minimisation du *NFA* nécessite idéalement l'exploration combinatoire de tous les sous-domaines de l'image. Cette recherche exhaustive n'est cependant pas envisageable en pratique dans la mesure où, même pour une imagerie de 256 pixels, il faudrait analyser  $2^{256} \simeq 1,15 \cdot 10^{77}$  sous-domaines, alors que les images réelles contiennent plutôt de l'ordre de 10 000 pixels chacune. En remarquant que les sous-domaines pour lesquels l'estimation des moyennes est meilleure sont ceux pour lesquels les résidus quadratiques sont minimaux, nous proposons de restreindre l'exploration aux sous-domaines sélectionnés dans l'esprit de la stratégie RANSAC.

L'algorithme que nous adoptons pour la détection repose donc sur la collaboration entre le principe d'échantillonnage aléatoire que nous venons de décrire et le critère probabiliste du *nombre de fausses alarmes* défini dans le chapitre 7 de manière à s'affranchir du choix d'un seuil arbitraire de décision lors de la phase d'estimation des paramètres tout en réduisant le nombre de sous-domaines à tester. Une collaboration de ce type a déjà été mise en oeuvre

par [Moisan et Stival, 2004] pour la détection de rigidité et l'estimation de la matrice fondamentale en vision stéréoscopique. Cette méthode permet la détection de rigidité et l'estimation de la matrice fondamentale lorsque les données contiennent jusqu'à 90% de points aberrants, performances très supérieures à celles obtenues jusqu'à présent par les meilleures méthodes robustes dans un contexte similaire [Torr et Murray, 1997, Salvi et al., 2001]. Dans le cadre de notre étude, le *NFA* permet d'évaluer le degré de significativité d'un sous-ensemble relativement à une labellisation, et de décider ainsi si le sous-ensemble le plus significatif trouvé par l'algorithme peut être considéré comme significativement cohérent avec la labellisation ( $NFA \ll 1$ ) ou si ce niveau de cohérence aurait pu être obtenu par hasard.

Dans un premier temps, nous présentons section 8.2 les algorithmes de détection dans les cas monodimensionnel et multidimensionnel, *i.e.* selon le nombre d'images **BR** utilisées. La section 8.3 aborde le problème du nombre d'itérations à effectuer, en pratique, de manière à garantir une détection contenant un minimum de fausses alarmes. Pour finir, la section 8.4 présente une stratégie de sélection de sous-systèmes en vue d'optimiser l'algorithme.

## 8.2. Algorithme

L'algorithme que nous proposons pour minimiser le *NFA* repose sur une exploration aléatoire des sous-domaines de l'image **BR**, gouvernée par la stratégie d'échantillonnage aléatoire adoptée pour l'estimation des moyennes des classes (*cf.* 8.1). Même si les approches multidimensionnelle et monodimensionnelles sont très proches, le fait qu'un pixel donné ne soit pas nécessairement valide à toutes les dates, dans le cas multidimensionnel, nécessite une approche particulière. Nous présentons donc, dans la section suivante, l'algorithme monodimensionnel avant d'aborder certains aspects spécifiques au multidimensionnel dans la section 8.2.2.

### 8.2.1. Cas monodimensionnel

Avant de rechercher le domaine qui minimise le *nombre de fausses alarmes*, le calcul du *NFA* nécessite l'estimation des moyennes des classes. Dans l'esprit de la stratégie d'échantillonnage aléatoire évoquée section 8.1, nous proposons, à chaque itération, de sélectionner  $L$  pixels du domaine image, et d'estimer les moyennes  $\mu$  à partir des valeurs de ces  $L$  pixels. Le système à résoudre, alors bien déterminé, est formé de  $L$  équations de type

$$v(y) = \sum_l \alpha_l(y) \mu(l),$$

pour lesquelles les observations  $v(y)$  et les proportions de chaque label dans le pixel  $y$ ,  $\alpha_l(y)$ , sont connues. Ce système étant déterminé à partir du tirage aléatoire de  $L$  pixels dans l'image, rien ne garantit qu'il soit numériquement inversible (les  $L$  labels ne sont pas nécessairement représentés par ces pixels). Dans l'objectif de pouvoir traiter aussi les systèmes non-inversibles, nous utilisons la méthode de décomposition en valeurs singulières (SVD *cf.* [Press et al., 1988]) pour estimer la

famille  $\{\mu(l)\}_{l \in \mathcal{L}}$  qui correspond à l'observation des  $L$  pixels. Cependant, nous verrons Section 8.4 que la restriction de l'étude aux systèmes inversibles (et bien conditionnés) permet d'accélérer le temps de calcul, les systèmes non inversibles ne conduisant généralement pas au  $NFA$  minimum. Une méthode de résolution telle que la décomposition LU est alors plus indiquée.

La définition du  $NFA$  repose sur la probabilité, pour un sous-domaine  $D$ , d'observer une erreur quadratique particulièrement faible sous l'hypothèse  $H_0$  (modèle naïf). Une fois cette famille de moyennes estimée, l'erreur quadratique obtenue entre l'observation et l'estimation est donc calculée en chaque pixel  $y$  du domaine, soit

$$r(y) = \left( v(y) - \sum_{l \in \mathcal{L}} \alpha_l(y) \mu(l) \right)^2.$$

Le calcul du  $NFA$  associé à un sous-domaine  $D$  se fait alors directement à partir des paramètres liés aux images entrées et de la somme des erreurs quadratiques obtenues pour tous les pixels de  $D$ . Cependant, la proposition suivante permet d'optimiser la recherche du  $NFA$  minimum en remarquant que, pour une image donnée, le  $NFA$  ne dépend pas directement d'un sous-domaine  $D$  mais seulement de son cardinal et, de manière monotone, des erreurs quadratiques cumulées sur le sous-domaine  $D$  ( $\sum_{y \in D} r(y)$ ).

**Proposition 8.2.1** *Notons, pour  $D \subset \mathcal{D}_{BR}$ ,  $\delta^2(v_D) = \sum_{y \in D} r(y)$ . Si  $\mathcal{D}_{BR} = \{y_i\}_{i=1 \dots |\mathcal{D}_{BR}|}$  et pour tout  $(y_i, y_j) \in \mathcal{D}_{BR}^2$ , on a  $i \leq j \Leftrightarrow r(y_i) \leq r(y_j)$ , alors*

$$\arg \min_{\substack{D \subset \mathcal{D}_{BR} \\ |D|=q}} NFA(|D|, \delta(v_D), \sigma) = \{y_1, \dots, y_q\}.$$

**Démonstration.**

Si  $r$  est ordonné alors pour tout  $D \in \mathcal{D}_{BR}$ ,  $\delta^2(v_D) \geq \delta^2(v_{\{y_1, y_2, \dots, y_{|D|}\}})$ . La fonction  $NFA$  étant croissante en  $\delta$ , son minimum est obtenu pour le sous-domaine  $\{y_1, \dots, y_q\}$ .  $\square$

En triant les résidus quadratiques par ordre croissant, il suffit donc de considérer chaque sous-domaine dans l'ordre de ses résidus pour minimiser le  $NFA$ . En effet, rappelons que la définition du  $NFA$  associé à un sous-domaine  $D$  repose sur la probabilité que le résidu cumulé observé soit particulièrement faible pour l'hypothèse *a contrario* ( $H_0$ ), *i.e.* inférieure à un certain  $\delta^2$ , et si  $\delta_1^2 < \delta_2^2$ , alors la probabilité que le résidu cumulé soit inférieur à  $\delta_1^2$  est inférieure à celle qu'il soit inférieur à  $\delta_2^2$ .

L'itération de ce processus un grand nombre de fois permet une exploration aléatoire des sous-domaines, et la recherche du  $NFA$  minimum se fait sur toutes ces itérations. Le sous-domaine qui minimise le  $NFA$  est alors celui qui est le plus cohérent avec la labellisation, *i.e.* celui qui rejette le mieux l'hypothèse naïve. L'algorithme 3 utilise, en entrée, une labellisation **HR**, une image **BR**, et admet pour unique paramètre le nombre total d'itérations. Il permet d'obtenir, en sortie, les moyennes des classes associées au  $NFA$  minimal obtenu, la valeur du  $NFA$  minimal et le sous-domaine  $D \subset \mathcal{D}_{BR}$  correspondant au  $NFA$  minimal, *i.e.* pour lequel l'image est cohérente avec la labellisation.

- Assigner  $\sigma$  à l'écart-type de l'image.
- Initialiser  $\delta_{min}^2$  et  $NFA_{min}$  à  $+\infty$ .
- Itérer  $N$  fois
  1. tirer uniformément et indépendamment  $L = |\mathcal{L}|$  pixels  $y$  de  $\mathcal{D}_{BR}$ ,
  2. estimer la famille  $\mu$  des moyennes des labels à l'aide des équations
 
$$v(y, \mu) = \sum_{l \in \mathcal{L}} \alpha_l(y) \mu(l),$$
 formées pour ces  $L$  pixels,
  3. calculer l'erreur quadratique  $r(y) = (v(y) - \sum_{l \in \mathcal{L}} \alpha_l(y) \mu(l))^2$ , pour tout  $y \in \mathcal{D}_{BR}$ ,
  4. trier  $\mathcal{D}_{BR}$  en un vecteur  $(y_i)_{1 \leq i \leq |\mathcal{D}_{BR}|}$  selon l'ordre croissant des erreurs obtenues  $r(y_i)$  : les  $L$  premières au moins sont nulles.
  5. initialiser  $\delta^2 = \sum_{i=0}^L r(y_i)$  (i.e.  $\delta^2 = 0$ ),
  6. pour chaque indice  $i = L + 1$  à  $i = |\mathcal{D}_{BR}|$ ,
    - ajouter à  $\delta^2$  l'erreur résiduelle  $r(y_i)$  du pixel  $y_i$ ,
    - si c'est la plus petite valeur obtenue sur toutes les itérations, à taille de domaine testé égale ( si  $\delta^2 < \delta_{min}^2[i]$ ), alors
      - mettre à jour  $\delta_{min}^2[i]$ ,
      - calculer la valeur du  $NFA$  correspondante,
      - si ce  $NFA$  est le meilleur obtenu jusqu'à présent pour un domaine de cette taille, (si  $NFA[i] \leq NFA_{min}$ ) alors
        - mettre à jour  $NFA_{min}$  et sauvegarder l'ensemble  $D = \{y_k\}_{k=1..i}$ .
    - fin si
  7. fin pour
- fin itérer

**Algorithme 3:** Minimisation du  $NFA$  dans le cas monodimensionnel.

### 8.2.2. Cas multidimensionnel

L'expression multidimensionnelle du *nombre de fausses alarmes* (7.4.53) est analogue à celle obtenue dans le cas monodimensionnel. Cependant, il n'est pas envisageable d'adopter une approche séquentielle (comparaison des *NFA* minimaux obtenus séparément avec l'algorithme 3 pour chaque image) pour détecter les changements sur une séquence d'images. En effet, le sous-domaine spatial correspondant à l'argument du *NFA* minimal calculé séparément pour chaque date n'est pas révélateur des changements qui apparaissent sur les autres images de la séquence. Ce type d'approche pourrait davantage être utilisé dans le but de déterminer quelle image de la séquence est la plus cohérente avec la labellisation de référence que pour la détection de changements dans un cadre spatio-temporel, où des changements sont susceptibles de se produire à différentes dates sans être visibles sur les autres acquisitions. Toutefois, un raisonnement sur l'intersection des domaines les plus significatifs obtenus séparément à chaque date pourrait être envisagé, bien que lourd à mettre en oeuvre lorsque l'ensemble des pixels valides est différent à chaque date. De plus, dans le contexte de la détection de changements de type d'occupation du sol, il est nécessaire de considérer l'évolution temporelle des intensités et, donc, d'analyser la séquence multitemporelle dans son ensemble plutôt que séparément pour chaque date. Par conséquent, nous proposons d'utiliser l'expression multidimensionnelle du *NFA* dans le cadre d'une approche vectorielle pour détecter le domaine spatio-temporel des changements. Plus précisément, il s'agit de considérer chaque point de la grille spatiale de manière vectorielle, et de minimiser le *NFA* à partir des différences vectorielles. Dans ce cas, si un pixel est *incorrect* sur une image, il est rejeté pour toute la série.

L'approche vectorielle peut aussi être utilisée, par exemple, pour les images multispectrales où chaque image admet le même domaine de définition. Dans le cas multitemporel, pour une date considérée, la valeur observée n'est généralement pas utilisable ou accessible pour tous les pixels du domaine. De plus, l'ensemble des pixels pour lesquels les valeurs sont aberrantes ou manquantes varie selon les dates d'acquisitions. L'approche vectorielle ne peut donc pas être mise en oeuvre directement à partir de l'algorithme 3 dans le cas de séquences temporelles, à moins de restreindre l'étude au domaine spatial sur lequel les valeurs acquises sont valides à toutes les dates. Le domaine d'analyse risque alors d'être considérablement réduit. En revanche, nous proposons (cf. algorithme 4) une variante de l'algorithme 3 adaptée à l'utilisation d'un domaine spatio-temporel restreint aux pixels corrects, domaine variable selon les dates et pixels considérés.

Dans cet objectif, rappelons les notations de la section 7.4 pour définir les domaines de validité. Si  $\mathcal{T}$  représente l'ensemble des dates d'acquisition des images de la séquence, et  $\mathcal{D}_{BR}$  le domaine image, le domaine spatio-temporel sur lequel les données sont valides est alors un sous-domaine de  $\mathcal{D}_{BR} \times \mathcal{T}$ , noté  $\Omega$ . Notre objectif est alors de détecter dans quel sous-domaine spatio-temporel  $\omega$  du domaine valide  $\Omega$  la séquence d'images **BR** est encore cohérente avec la labellisation haute résolution. L'algorithme est très proche de l'algorithme 3 décrit pour le cas monodimensionnel, avec quelques différences spécifiques principalement dûes au fait que les domaines spatiaux valides ne soient pas identiques à chaque date. Pour tout pixel  $y \in \mathcal{D}_{BR}$ , notons

$\mathcal{T}_y$  l'ensemble de toutes les dates auxquelles le pixel  $y$  est valide. De même que dans le cas monodimensionnel, l'algorithme repose sur la sélection aléatoire de  $|\mathcal{L}| \times |\mathcal{T}|$  pixels dans le domaine spatio-temporel  $\Omega$  en vue d'estimer les moyennes des classes  $(\mu_t(l))_{l,t}$ . Cette estimation se fait par la résolution du système bien déterminé, extrait à partir des pixels sélectionnés, de la forme

$$v(y, t) = \sum_{l \in \mathcal{L}} \alpha_l(y) \mu_t(l).$$

La sélection aléatoire des pixels doit donc assurer la disponibilité de l'information à chaque date pour permettre l'estimation des moyennes, *i.e.* il ne s'agit pas de tirer les pixels au sein du domaine  $\mathcal{D}_{\text{BR}} \times \mathcal{T}$  mais dans le sous-domaine des pixels valides  $\Omega$ . Les pixels sélectionnés sont alors susceptibles d'être différents à chaque date. De plus, comme dans le cas monodimensionnel, nous utilisons la méthode de décomposition en valeurs singulières pour résoudre le système extrait en vue de permettre l'étude de tous les systèmes sélectionnés, même lorsqu'ils ne sont pas inversibles. Toutefois, nous verrons dans la section 8.4 qu'une telle exploration n'est pas nécessaire et que l'étude peut raisonnablement être restreinte aux cas des systèmes inversibles, une décomposition de type LU (*cf.* [Press et al., 1988]) sera alors adoptée. Une fois les moyennes estimées, les résidus quadratiques obtenus entre la séquence d'images observée et sa reconstruction (d'après le modèle linéaire de mélange 7.4.50 et les moyennes estimées) sont calculés pour tout pixel spatio-temporel  $(y, t)$  du domaine valide  $\Omega$ , soit

$$r(y, t) = \left( v_t(y) - \sum_{l \in \mathcal{L}} \alpha_l(y) \mu_t(l) \right)^2.$$

Une approche analogue au cas monodimensionnel serait de considérer les sous-domaines  $\omega \subset \Omega$  de manière incrémentale, à travers les résidus cumulés sur les sous-domaines spatio-temporels  $\omega$ . Une telle approche pourrait conduire à rejeter un pixel  $y$  à cause de l'erreur mesurée en un couple  $(y, t)$  alors que, en dehors de cette date  $t$ , le pixel  $y$  est toujours cohérent avec la labellisation. Il nous semble important, au regard de l'application, de ne pas considérer les composantes temporelles séparément les unes des autres pour un pixel donné de manière à ne pas rejeter systématiquement un pixel lorsque la valeur acquise pour une seule date n'est pas cohérente avec la labellisation. Par conséquent, plutôt que de considérer l'ensemble des couples  $(y, t)$  du domaine valide  $\Omega$ , nous proposons d'explorer chaque pixel  $y \in \mathcal{D}_{\text{BR}}$  directement pour toutes les dates valides  $t \in \mathcal{T}_y$ . Cette exploration revient à restreindre l'étude, pour tout pixel  $y \in D$ , à des paliers de taille  $|\mathcal{T}_y|$ . Dans cet objectif, nous décidons, plutôt que d'étudier directement les résidus quadratiques pour chaque couple  $(y, t) \in \Omega$ , de considérer un résidu quadratique moyen, noté  $\bar{r}$ , défini pour tout  $y \in \mathcal{D}_{\text{BR}}$  par

$$\bar{r}(y) = \frac{\sum_{t \in \mathcal{T}_y} r(y, t)}{|\mathcal{T}_y|}. \quad (8.2.1)$$

Finalement, il s'agit d'approximer le résidu  $r(y, t)$ , pour tout  $(y, t) \in \Omega$ , par le résidu moyen  $\bar{r}(y, t)$  et les sous-domaines spatio-temporels explorés sont alors de la forme  $\Omega_D = \cup_{y \in D} \{y\} \times \mathcal{T}_y$ . Rappelons que le *NFA* dépend, en dehors des paramètres directement liés aux observations,

uniquement du cardinal du sous-domaine valide  $\Omega_D$  et, de manière monotone, des résidus cumulés sur ce sous-domaine. La proposition suivante correspond à l'adaptation de la proposition 8.2.1 au cas multitemporel pour lequel l'exploration des sous-domaines est limitée à l'ensemble des sous-domaines valides de la forme  $\Omega_D$ , i.e. sous la contrainte que toutes les dates valides associées à chaque pixel considéré soient prises en compte. Pour tout sous-domaine spatio-temporel  $\Omega_D \in \Omega_{\mathcal{D}_{BR}}$ , le résidu cumulé ne dépend finalement que de son cardinal  $|\Omega_D|$ . La recherche du *NFA* minimum sur tous les sous-domaines spatio-temporels  $\Omega_D$  peut donc être optimisée en utilisant le résultat suivant.

**Proposition 8.2.2** *Soit  $\Omega_D$  le sous-domaine spatio-temporel ( $\Omega_D \subset \Omega_{\mathcal{D}_{BR}}$ ) défini pour tout  $D \subset \mathcal{D}_{BR}$  par  $\cup_{y \in D} \{y\} \times \mathcal{T}_y$ . Si  $\delta^2(v_{\Omega_D}) = \sum_{y \in D} \bar{r}(y) \times |\mathcal{T}_y|$  et, pour tout  $(y_i, y_j) \in \mathcal{D}_{BR}^2$ , si on a  $i \leq j \Leftrightarrow \bar{r}(y_i) \leq \bar{r}(y_j)$ , alors*

$$\arg \min_{\substack{D \subset \mathcal{D}_{BR} \\ |D|=q}} NFA(|\Omega_D|, \delta(v_{\Omega_D}), \sigma) = \cup_{i=1}^q \{y_i\} \times \mathcal{T}_{y_i}.$$

**Démonstration.**

Le vecteur  $\bar{r}$  est trié selon l'ordre croissant donc, pour tout  $D \subset \mathcal{D}_{BR}$ , si  $q = |D|$ ,

$$\sum_{y \in D} \bar{r}(y) \geq \sum_{i=1}^q \bar{r}(y_i),$$

et

$$\delta^2(v_{\Omega_D}) = \sum_{y \in D} |\mathcal{T}_y| \bar{r}(y) \geq \sum_{i=1}^q |\mathcal{T}_{y_i}| \bar{r}(y_i) = \delta^2(v_{\Omega_{\{y_1, \dots, y_q\}}}) \quad (8.2.2)$$

□

Dans ce but, nous considérons pour chaque pixel  $y$  du domaine spatial  $\mathcal{D}_{BR}$ , la moyenne  $\bar{r}(y)$  des erreurs quadratiques obtenues pour toutes les dates valides associées au pixel  $y$  ( $t \in \mathcal{T}_y$ ). Les erreurs moyennes sont ensuite triées par ordre croissant, et la recherche du domaine qui minimise le *NFA* est alors identique à celui adopté pour le cas monodimensionnel, en considérant les domaines spatio-temporels valides à chaque itération (cf. Algorithme 4).

Dans la section 8.2.3, nous présentons un exemple d'application de l'algorithme 3 sur un jeu de données simulées à titre d'illustration avant de poursuivre l'étude de quelques aspects numériques tels que le nombre d'itérations nécessaire à la convergence, section 8.3.

### 8.2.3. Exemple d'application monotemporelle

La labellisation de référence, présentée figure 8.1 (a), a été simulée d'après des caractéristiques provenant d'extraits d'images réelles, comme décrit section 10.2. Une valeur moyenne  $\mu(l)$  est attribuée à chaque label  $l$ , arbitrairement dans  $\{0, 0.1, \dots, 0.8\}$ . L'image **HR** correspondant à la labellisation est alors simulée en ajoutant à l'image ainsi obtenue un bruit blanc de moyenne nulle

- Assigner  $\sigma$  à l'écart-type de l'image.
- Initialiser  $\delta_{min}^2$  et  $NFA_{min}$  à  $+\infty$ .
- Itérer  $N$  fois
  1. tirer uniformément et indépendamment  $|\mathcal{L}| \times |\mathcal{T}|$  pixels  $(y, t)$  de  $\Omega$ ,
  2. pour chaque date  $t$ , estimer la famille  $\mu$  des moyennes des labels à l'aide des équations
 
$$v_t(y, \mu) = \sum_{l \in \mathcal{L}} \alpha_l(y) \mu_t(l),$$
 formées pour ces  $|\mathcal{L}| \times |\mathcal{T}|$  pixels,
  3. calculer l'erreur quadratique,  $\forall (y, t) \in \Omega$ ,
 
$$r(y, t) = (v_t(y) - \sum_{l \in \mathcal{L}} \alpha_l(y) \mu_t(l))^2$$
  4. calculer l'erreur moyenne sur l'ensemble des dates valides :  $\forall y \in \Omega_{\mathcal{D}_{BR}}$ ,
 
$$\bar{r}(y) = \frac{\sum_{t \in \mathcal{T}_y} r(y, t)}{|\mathcal{T}_y|}$$
  5. trier  $\mathcal{D}_{BR}$  en un vecteur  $(y_i)_{1 \leq i \leq |\mathcal{D}_{BR}|}$  selon l'ordre croissant des erreurs moyennes  $\bar{r}(y_i)$ .
  6. initialiser  $\delta^2 = \sum_{i=1}^L \bar{r}(y_i) \times |\mathcal{T}_{y_i}|$  (i.e.  $\delta^2 = 0$ ),
  7. pour chaque indice  $i = L + 1$  à  $i = |\mathcal{D}_{BR}|$ ,
    - ajouter à  $\delta^2$  l'erreur résiduelle  $\bar{r}(y_i) \times |\mathcal{T}_{y_i}|$
    - si c'est la plus petite valeur obtenue sur toutes les itérations, à taille de domaine testé égale ( si  $\delta^2 < \delta_{min}^2[i]$ ), alors
      - mettre à jour  $\delta_{min}^2[i]$ ,
      - calculer la valeur du  $NFA$  correspondante,
      - si ce  $NFA$  est le meilleur obtenu jusqu'à présent pour un domaine de cette taille, (si  $NFA[i] \leq NFA_{min}$ ) alors
        - mettre à jour  $NFA_{min}$  et sauvegarder l'ensemble  $\Omega_D = \cup_{k=1}^i \{y_k\} \times |\mathcal{T}_{y_k}|$ .
      - fin si
    - fin si
  8. fin pour
- fin itérer

**Algorithme 4:** Minimisation du  $NFA$  dans le cas multitemporel.



et d'écart-type  $\sigma_b = 0.05$  de manière à modéliser la variabilité intra-classes. A partir de cette image, nous simulons une autre image contenant deux changements de type différents (cf. 8.1 (b)) : dans le coin en haut à gauche, une classe est scindée en deux classes d'intensités moyennes très proches et, en bas à droite, plusieurs parcelles de labels différents sont recouvertes partiellement d'une zone circulaire de bruit blanc centré d'écart-type 0.1 (simulation d'une zone inondée ou brûlée par exemple). Même si cette forme circulaire est naïve pour simuler des changements naturels, elle n'a pas d'importance puisque l'information spatiale ne joue aucun rôle au sein de l'approche que nous proposons. Pour finir, cette image moyennée par blocs de taille  $16 \times 16$  est utilisée comme entrée de l'algorithme (figure 8.2). L'objectif est alors de détecter les changements apparus dans l'image **BR** par rapport à la labellisation **HR**.

L'algorithme est capable, à partir de la labellisation **HR** et de l'image **BR** (figure 8.1 et 8.2), de détecter les 27 pixels représentés en noir figure 8.2. Les lignes blanches superposées aux sous-domaines détectés correspondent à la trace des changements introduits à haute résolution, elles permettent d'observer visuellement que les pixels détectés par l'algorithme sont bien ceux qui ont été touchés par des changements. Tous les pixels touchés par le disque de changements ont été détectés, par contre, certains pixels contenant une très faible proportion de changements n'ont pas été détectés par la méthode. Plus précisément, la parcelle qui se distingue du fond, en haut à gauche de l'image (figure 8.1), intersecte 8 pixels basse résolution, dans une proportion allant de 1.5% à 98%. Seuls six pixels ont été détectés. les deux restant correspondent au deux plus petites valeurs de proportion, i.e. 1.5 et 5.4%. Notons que la valeur moyenne simulée pour cette parcelle est particulièrement proche de celle du fond. La même expérience a été réalisée lorsque la valeur moyenne attribuée à la parcelle est légèrement supérieure. L'algorithme détecte alors 7 pixels sur 8, seul le pixel intersectant 1.5% de la parcelle n'a alors pu être détecté.

Le graphe (figure 8.2) présente, en fonction de la taille  $k$  des sous-domaines considérés, l'erreur quadratique minimale obtenue sur tous les sous-domaines de cardinal  $k$  (croix rouges) et la valeur du  $NFA$  minimale obtenue pour tous les sous-domaines de cardinal  $k$  (pointillés verts) (en échelle  $\log_{10}$ ). La valeur minimale du  $NFA$  est obtenue précisément en 229, indiquant que le sous-domaine cohérent avec la labellisation est de cardinal 229, soit 27 pixels de changements ( $256 - 229 = 27$ ). Remarquons que la courbe de l'erreur quadratique minimale à elle seule ne présente pas de point singulier, elle permettrait donc difficilement une telle précision dans la détection.

Avant d'analyser plus précisément les performances du modèle de détection, nous discutons, dans les sections 8.3 et 8.4, deux points importants pour sa mise en oeuvre : le nombre d'itérations nécessaire à la convergence de l'algorithme et la stratégie de sélection de sous-systèmes pour l'estimation des moyennes des classes.

### 8.3. Nombre d'itérations nécessaire

La méthode de détection est fondée, numériquement, sur une stratégie d'échantillonnage aléatoire qui, pour bien fonctionner, doit être répétée un grand nombre de fois. Une estimation du

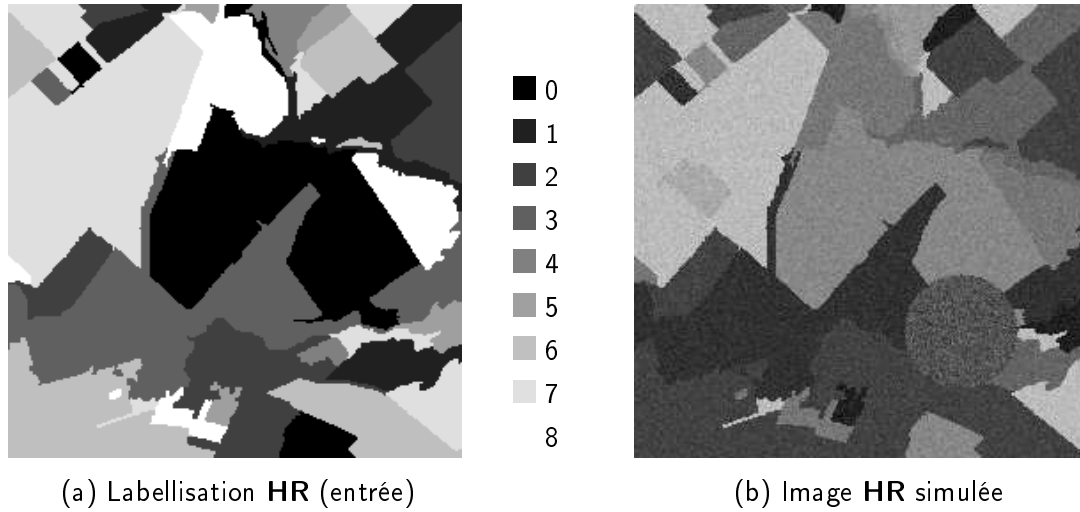


FIG. 8.1.: A gauche, la labellisation **HR** contenant 9 labels (entrée de l'algorithme). A droite, image **HR** simulée à partir de la labellisation en introduisant 2 zones de changements de type différents.

nombre d'itérations nécessaires pourrait être envisagée de la même manière que dans les travaux de [Moisan et Stival, 2004] sur la détection de rigidités. En effet, supposons que l'image **BR** contienne  $n$  pixels, la labellisation  $L$  labels, et que sur ces  $n$  pixels,  $k$  soient cohérents avec la labellisation (pixels dit *corrects*). Il y a donc  $n - k$  pixels de changements (*incorrects*) dans l'image. La valeur du *NFA* associée à un sous-domaine  $D$  admet pour paramètre l'erreur quadratique cumulée sur le sous-domaine  $D$ , erreur dont la valeur dépend de la qualité de l'estimation des moyennes des classes. La valeur du *NFA* dépend donc des  $L$  pixels sélectionnés, et la détection est d'autant plus performante qu'ils correspondent bien à la labellisation. Le nombre d'itérations nécessaire peut alors être considéré comme lié à la probabilité de sélectionner  $L$  pixels *corrects* dans l'image ( $= \frac{C_k^L}{C_n^L}$ ). Plus précisément, si  $q$  représente la probabilité de sélectionner au moins un sous-ensemble de  $L$  pixels *corrects* en effectuant  $N$  tirages,

$$q = 1 - \mathbb{P}(\text{"ne pas tirer } L \text{ pixels corrects"}) \quad (8.3.3)$$

$$= 1 - \left(1 - \frac{C_k^L}{C_n^L}\right)^N. \quad (8.3.4)$$

A partir de là, nous pouvons aisément calculer le nombre  $N$  de tirages nécessaires pour avoir  $q$  de chances de sélectionner au moins un bon sous-ensemble de  $L$  pixels dans l'image, soit

$$N = \frac{\ln(1 - q)}{\ln\left(1 - \frac{C_k^L}{C_n^L}\right)}. \quad (8.3.5)$$

Cependant, dans notre contexte, l'estimation des moyennes des classes passe par l'inversion (ou la pseudo-inversion) d'une matrice extraite de la matrice des proportions de chaque classe dans chaque pixel (matrice extraite de taille  $L \times L$ ). Si les pixels sélectionnés ne représentent pas,

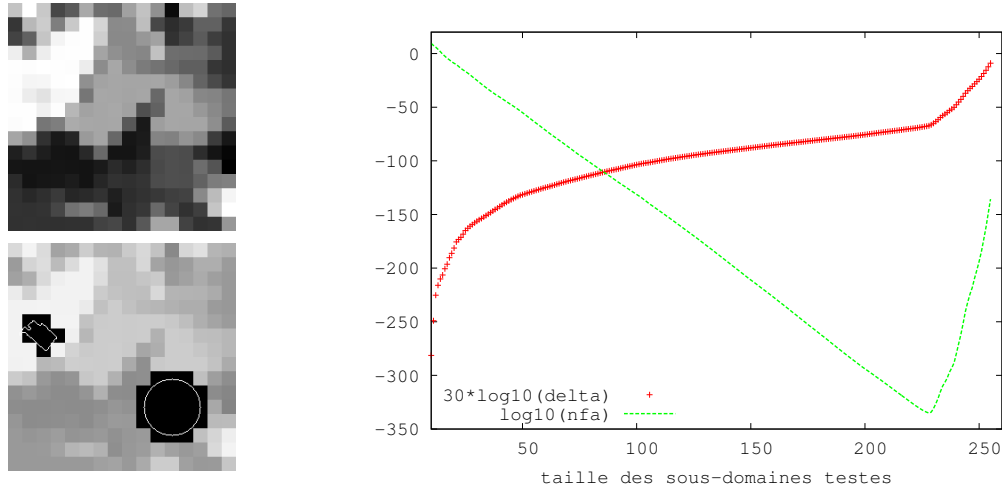


FIG. 8.2.: A gauche, en haut : image **BR** simulée par moyennage de l'image figure 8.1(b) par blocs de taille  $16 \times 16$ ; en bas, les pixels de changements détectés (27) sont représentés en noir, et la trace des contours des changements introduits à haute résolution est superposée en blanc (la taille des pixels **BR** a été agrandie pour plus de lisibilité). A droite, les croix rouges représentent l'erreur quadratique minimale obtenue en fonction de la taille du sous-domaine  $D$  étudié (échelle  $\log_{10}$ ); en vert, les pointillés représentent la valeur minimale du *NFA* obtenue en fonction de la taille du sous-domaine considéré.

à eux tous, les  $L$  classes représentées dans la labellisation, alors la matrice extraite est de rang strictement inférieur à  $L$  et elle ne permet pas d'estimer les moyennes des classes. Finalement, il faut donc tirer au moins une fois  $L$  pixels corrects et représentant les  $L$  labels. La probabilité de sélectionner  $L$  pixels à la fois corrects et représentant  $L$  labels est plus compliquée à calculer dans la mesure où chaque pixel **BR** est susceptible de représenter plusieurs labels. A titre indicatif, supposons que tous les pixels soient *purs* (i.e. chaque pixel représente un label unique) et calculons la probabilité de sélectionner  $L$  pixels corrects représentant les  $L$  labels. Si, pour  $i = 1 \dots L$ ,  $k_i$  pixels représentent le label  $i$  et seuls  $k$  pixels de l'image sont corrects au total, alors l'hypothèse des pixels *purs* permet d'avoir  $k = k_1 + \dots + k_L$  et la probabilité de sélectionner  $L$  pixels corrects et représentant  $L$  labels vaut

$$\frac{C_{k_1}^1 C_{k_2}^1 \dots C_{k_L}^1}{C_n^L}. \quad (8.3.6)$$

La probabilité  $q'$  de sélectionner au moins un sous-ensemble de  $L$  pixels corrects et représentant les  $L$  labels vaut alors

$$q' = 1 - \left( 1 - \frac{C_{k_1}^1 C_{k_2}^1 \dots C_{k_L}^1}{C_n^L} \right)^N. \quad (8.3.7)$$

Pour avoir une idée de l'ordre de grandeur du nombre d'itérations dans les deux cas, considérons le cas des données simulées utilisées dans le chapitre 10. L'image **BR** contient  $n = 256$  pixels et la labellisation 10 labels. Supposons que 200 pixels soient corrects, le nombre de tirages nécessaires pour avoir, par exemple, 95% de chances de tirer au moins une fois un sous-ensemble de  $L$  pixels

corrects vaut alors 35 environ, d'après (8.3.5). Maintenant si nous supposons, en plus, que les labels sont équidistribués, le nombre d'itérations nécessaire pour tirer  $L$  pixels corrects et représentants les  $L$  labels vaut alors, d'après (8.3.7)

$$N = \frac{\ln(1 - q')}{\ln\left(1 - \frac{(\frac{k}{L})^L}{C_n^L}\right)} \quad (8.3.8)$$

$$\simeq 6908. \quad (8.3.9)$$

Ces hypothèses, très simplificatrices, ne sont pas réalistes en général. En particulier, certaines classes, telles que l'eau, sont vouées à être minoritaires dans la plupart des cas. De plus, avec les rapports de résolution tels que ceux que nous considérons, chaque pixel **BR** représente en général un mélange de plusieurs classes. Cette estimation n'est donc pas utilisable en pratique. Finalement, le nombre d'itérations est donc fixé de manière expérimentale.

Pour un jeu de données simulées comme décrit section 10.3.1, avec 30% de bruit impulsif, la figure 8.3 représente, en fonction du nombre d'itérations  $N$ , l'erreur quadratique obtenue entre les moyennes associées au *NFA* minimal pour  $N$  itérations et les moyennes solutions (*i.e.* moyennes associées au *NFA* minimal pour 1 000 000 d'itérations). Cette figure permet d'observer la convergence de l'algorithme dès 100 000 itérations pour ce jeu de données.

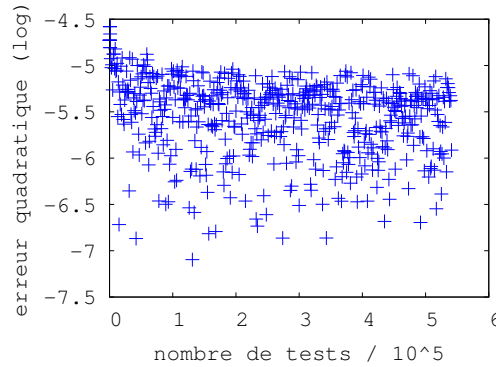


FIG. 8.3.: Erreur quadratique (en échelle log) obtenue pour l'estimation des moyennes en fonction du nombre d'itérations réalisées ( $10^5$ ). Résultat obtenu avec tirage aléatoire des pixels.

Dans le cadre de notre étude, la stratégie adoptée pour l'échantillonnage aléatoire en lui-même est importante pour la performance de l'algorithme dans la mesure où la sélection aléatoire de pixels dans le domaine image conduit souvent à des systèmes non inversibles, dans lesquels tous les classes ne sont pas nécessairement représentées. De nombreuses itérations sont alors amorcées inutilement, avant d'observer que le système n'est pas inversible. Pour éviter trop de calculs inutiles, l'objectif de la section 8.4 est d'étudier les sous-systèmes susceptibles d'être intéressants pour l'estimation des moyennes, en vue d'élaborer une stratégie de sélection capable d'optimiser la méthode de détection.

## 8.4. Sélection de sous-systèmes

Pour détecter le sous-domaine le plus cohérent avec une labellisation contenant  $L$  labels, la méthode décrite section 8.2 est basée sur la sélection aléatoire, dans le cas monodimensionnel, de  $L$  pixels du domaine  $\mathcal{D}_{BR}$ . Dans le cas multidimensionnel ( $\mathcal{D}_{BR} \times T$ ), la problématique est la même mais il s'agit de sélectionner  $L \times T$  pixels. Pour simplifier les notations, nous considérons uniquement le cas monodimensionnel dans cette section. A partir des  $L$  pixels sélectionnés,  $(y_1, \dots, y_L)$ , un sous-système de  $L$  équations à  $L$  inconnues défini par

$$\left( v(y_i) = \sum_{l \in L} \alpha_l(y_i) \mu(l) \right)_{i=1 \dots L}$$

est extrait du système qui pourrait être formé à partir de toute l'image  $\mathcal{D}_{BR}$ . Ce système extrait peut s'écrire aussi sous la forme matricielle  $Ax = b$ , où  $A = (\alpha_l(y_i))_{i,l}$  représente la matrice des proportions,  $x = (\mu_l)_l$  le vecteur des moyennes des classes et  $b = (v(y_i))_i$  le vecteur des observations associées aux  $L$  pixels sélectionnés.

En pratique, pour qu'une itération soit utile, il faut que les moyennes associées à chaque classe puissent être estimées. En tirant  $L$  pixels aléatoirement (uniformément entre 0 et  $|\mathcal{D}_{BR}|$ ), il est possible que certaines classes ne soient représentées par aucun pixel, le vecteur des moyennes ne peut alors pas être estimé correctement. Avant de chercher à élaborer une stratégie de sélection de sous-systèmes capable d'éviter ce type de cas, nous proposons d'analyser empiriquement la fréquence d'apparition de matrices non inversibles dans un contexte aléatoire. Dans cet objectif, nous considérons le conditionnement d'une matrice comme mesure d'*inversibilité*. Rappelons que le conditionnement d'une matrice carrée  $A$  est défini par

$$\text{cond}(A) = |||A||| \ |||A^{-1}|||,$$

où  $|||.|||$  est une norme matricielle subordonnée. Dans ce document, on considère en particulier les normes matricielles usuelles, subordonnées aux normes vectorielles  $\|\cdot\|_p$ , où  $p = 1, 2, \infty$ . Le conditionnement d'une matrice est toujours supérieur ou égal à 1 et, par convention, il est infini lorsque la matrice est singulière quelque soit la norme utilisée.

Les histogrammes de fréquence des conditionnements obtenus pour différents types de matrices aléatoires, de taille  $6 \times 6$  (100 000 tirages de matrices) sont présentés figure 8.4. La figure 8.4 (a) présente l'histogramme de fréquence des valeurs de conditionnement obtenues pour 100 000 matrices quelconques sélectionnées aléatoirement, uniformément dans  $[0, 1]$ . Remarquons que les valeurs de conditionnement obtenues sont toutes concentrées approximativement sur l'intervalle  $[0.8, 4]$  en échelle log, soit entre les valeurs 6 et 10 000. Ces matrices quelconques tirées aléatoirement (uniformément dans  $[0, 1]$ ) peuvent donc toutes être considérées comme numériquement inversibles. Cependant, les matrices sélectionnées pour notre étude possèdent des propriétés particulières puisqu'il s'agit de matrices de proportions. Elles sont donc, par définition, stochastiques en ligne (la somme en ligne vaut 1). De plus, remarquons que si  $A$  représente la proportion de chaque label dans un pixel basse résolution, son terme général est de la forme  $\frac{a}{N}$  où  $N$  est un

entier fixé représentant le nombre de pixels haute résolution contenus dans un pixel basse résolution, et  $a$  un entier compris entre 0 et  $N$ . La figure 8.4 (b) présente l'histogramme de fréquence de matrices de proportions simulées par tirage aléatoire tel que, pour tout  $(i, j)$ , le terme général  $a_{i,j} = a/N$ , avec  $a \in [0, N]$  (pour  $N = 15$ ). La plupart des matrices ainsi simulées sont aussi inversibles, même si quelques occurrences de conditionnement infini ont eu lieu (ici,  $1.e^{17}$  correspond à l'infini numérique). Pour finir, la figure 8.4 (c) présente l'histogramme obtenu pour des matrices de proportions réellement extraites à partir d'un jeu de données pour la détection de changements. Elle permet d'observer qu'environ 35% des matrices sélectionnées aléatoirement (uniformément) ne sont pas numériquement inversibles. Cette différence peut sembler surprenante mais il faut prendre en compte le fait que, pour une scène donnée, la fréquence d'apparition des valeurs de proportions n'est pas aléatoire : elle est directement liée à la structure de la labellisation relativement au domaine basse résolution, et certaines valeurs peuvent apparaître beaucoup plus fréquemment que d'autres. En pratique, la probabilité de sélectionner aléatoirement  $L$  pixels qui forment une matrice de proportions non-inversible est donc élevée, et implique autant d'itérations inutiles.

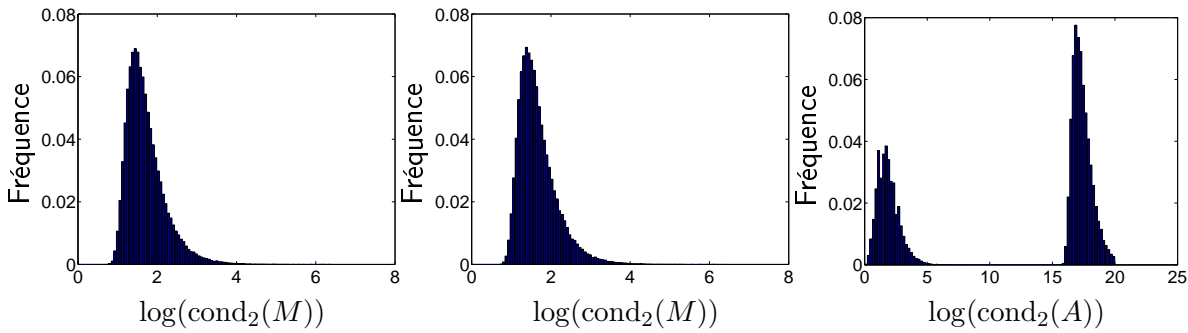


FIG. 8.4.: Histogrammes du conditionnement pour la norme 2 obtenu pour 100 000 matrices  $M$  de taille  $6 \times 6$  tirées aléatoirement (échelle log). Tirage uniforme de matrices quelconques dans  $[0, 1]$  figure (a), de matrices de proportions simulées figure (b) (terme général de la forme  $\frac{a}{15}$ ,  $a \in [0, 15]$ ), et de matrices de proportions extraites aléatoirement à partir d'un jeu de données figure (c).

A présent, vérifions que la valeur du meilleur  $NFA$  obtenu pour un jeu de données est bien liée à la valeur du conditionnement de la matrice utilisée pour l'estimation des moyennes. La figure 8.5 (a) présente les valeurs de  $NFA$  obtenues pour 100 000 itérations sur un jeu de données en fonction du conditionnement de la matrice  $A$  utilisée pour l'estimation des moyennes, dans le plan  $(\log(\text{cond}_2(A)), \log(NFA))$ . Remarquons que les plus petites valeurs de  $NFA$  ont bien été obtenues pour des valeurs faibles de conditionnement, même si des matrices singulières (conditionnement supérieur à  $1.e^{17}$ ) peuvent aussi permettre d'obtenir une faible valeur de  $NFA$ . Un conditionnement infini est synonyme d'instabilité mais il n'interdit pas l'obtention, par hasard, d'une solution correcte. Pour une analyse plus fine du cas des matrices inversibles, la figure 8.5 (b) présente la médiane, les 10ème et 90ème percentiles du  $NFA$  obtenu en fonction

du conditionnement de  $A$  pour le même échantillon restreint aux valeurs de conditionnement dont le logarithme est contenu dans l'intervalle  $[0, 6]$ , *i.e.* cas des matrices numériquement inversibles. Les valeurs du  $NFA$  sont d'autant plus grandes que le conditionnement est grand et ses valeurs minimales sont obtenues lorsque les matrices  $A$  sélectionnées sont de conditionnement inférieur à 3. L'histogramme bidimensionnel du conditionnement de  $A$  versus la valeur minimale du  $NFA$

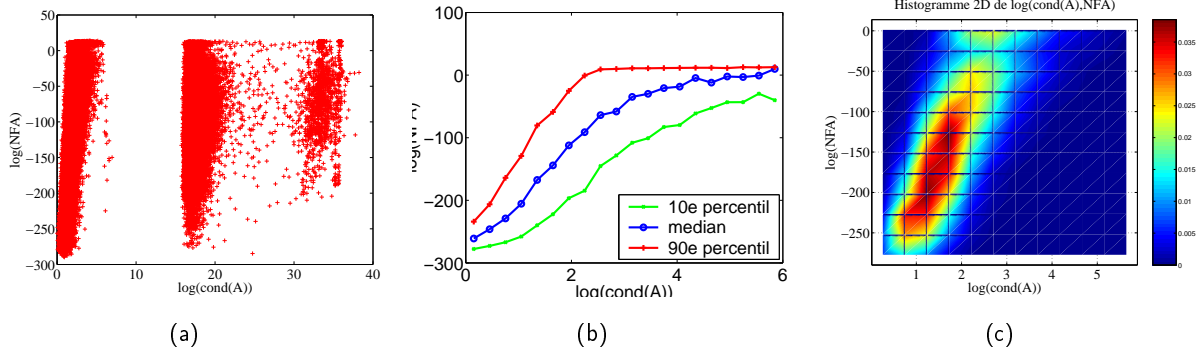


FIG. 8.5.: Valeur du  $NFA$  obtenu en fonction du conditionnement de la matrice  $A$  utilisée pour l'estimation des moyennes des classes, en échelle log. Résultats obtenus pour 100 000 itérations où les matrices sont sélectionnées aléatoirement (figure (a)). Figure (b) : valeurs médianes et des 10ème et 90ème percentiles obtenus sur la restriction de l'échantillon aux matrices numériquement inversibles, *i.e.* telles que  $\text{cond}_2(A) \leq 1.e^6$ . Figure (c) : l'histogramme de fréquence 2D du couple  $(\text{cond}_2(A), NFA)$ , en échelle log.

présenté figure 8.5 (c) permet de confirmer le lien étroit entre ces deux valeurs. Plus précisément, la figure 8.6 présente les histogrammes du  $NFA$  obtenus pour un intervalle de valeurs de conditionnement donné. Les matrices pour lesquelles le  $NFA$  est minimum semblent avoir, le plus souvent, un faible conditionnement. En effet, cette série d'histogrammes permet d'observer clairement que plus le conditionnement de la matrice sélectionnée est élevé, plus le pic de fréquence des valeurs de  $NFA$  est obtenu pour une grande valeur de  $NFA$ . De plus, remarquons notamment l'apparition d'un pic de fréquence des valeurs de  $NFA$  en 0 dès le 4ème histogramme, soit dès que  $\log(\text{cond}_2(A)) \geq 1.2156$ , *i.e.*  $\text{cond}_2(A) \geq 16.4$ . Cette remarque nous incite à penser que le fait de sélectionner les matrices de faible conditionnement permettrait de diminuer considérablement le nombre d'itérations pour un même résultat.

Pour améliorer la performance de l'algorithme de détection de changements et compte tenu des relations observées entre les valeurs minimales du  $NFA$  et la valeur des matrices exploitées pour l'estimation, nous proposons de rechercher un critère de sélection de matrices permettant le contrôle de son conditionnement et, par conséquent, de restreindre l'étude uniquement aux sous-systèmes suffisamment stables pour apporter une estimation correcte des moyennes des classes. En effet, les figures 8.5 et 8.6 montrent que toutes les itérations au cours desquelles des matrices non inversibles (ou à fort conditionnement) sont sélectionnées semblent peu utiles pour la minimisation du  $NFA$ .

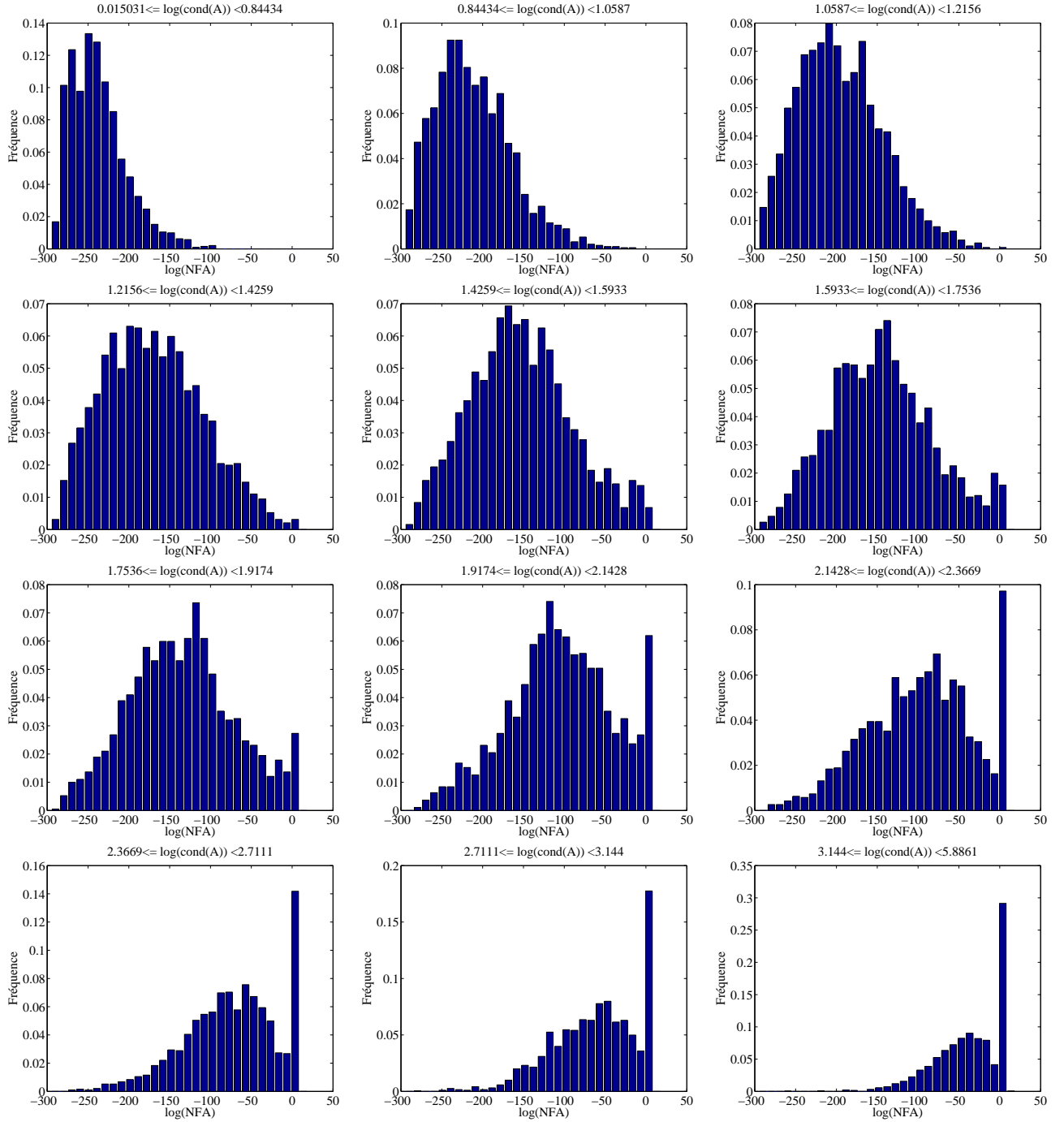


FIG. 8.6.: Histogrammes des valeurs de  $NFA$  obtenues pour un intervalle de valeurs du conditionnement de  $A$  donné. Chaque intervalle est défini de manière à compter le même nombre d'occurrences (1904). Plus le conditionnement de  $A$  est grand (en échelle  $\log$ ), plus les valeurs du  $\log(NFA)$  sont grandes.



Une première idée serait alors, à chaque itération, de tirer une matrice aléatoirement et de calculer son conditionnement avant de décider de poursuivre l'itération ou de répéter l'opération. Cependant, le calcul du conditionnement est trop coûteux pour être effectué à chaque itération, son temps de calcul étant supérieur au temps nécessaire à la résolution du système considéré (par décomposition en valeurs singulières ou décomposition LU). Remarquons que, pour être inversible, il est nécessaire que tous les labels soient représentés par l'ensemble des pixels sélectionnés. La première stratégie que nous proposons repose sur une condition intuitive de "poids minimal" sur les coefficients diagonaux (à une permutation près). Plus précisément, nous proposons de fixer une valeur seuil pour les proportions et de sélectionner une famille  $(y_1, \dots, y_L)$  telle que pour  $l = 1 \dots L$ , la proportion  $\alpha_l(y_l) \geq \text{seuil}$ . Avec cette stratégie, décrite par l'algorithme 5, toutes les classes sont représentées en proportion supérieure au seuil par au moins un pixel. Les valeurs

1. Pour  $l = 1$  à  $L$ 
  - trier le vecteur  $(\alpha_l(y_i))_{i=1 \dots |\mathcal{D}_{BR}|}$  par ordre croissant,
  - $I[l] = (y_i)_{i=1 \dots |\mathcal{D}_{BR}|}$  tel que  $\forall i \leq j, \alpha_l(y_i) \leq \alpha_l(y_j)$ ,
  - rechercher l'indice  $i_{\min}(l)$  tel que  $\forall i \geq i_{\min}(l), \alpha_l(y_i) \geq \text{seuil}$ ,
2. fin pour.
3. Pour  $l = 1$  à  $L$ 
  - tirer un entier  $m$  uniformément dans  $[i_{\min}(l), |\mathcal{D}_{BR}|]$ ,
  - le pixel correspondant  $y_l = I[l][m]$ ,
  - vérifier que le pixel n'a pas été sélectionné deux fois.
4. fin pour.

**Algorithme 5:** Sélection de sous-systèmes - stratégie 1.

de proportions qui correspondent à ces pixels sont stockées dans une matrice carrée, notée  $A$ , à partir de laquelle les moyennes des classes sont estimées avant de calculer le  $NFA$ . La figure 8.7 présente l'histogramme des valeurs de conditionnement obtenues pour 100 000 matrices sélectionnées aléatoirement (figure (a)) et par l'algorithme 5 avec un seuil fixé à 0.2 (figure (b)). Ces deux histogrammes sont étonnamment proches, et montrent que cette stratégie n'a pas d'effet sur le conditionnement des matrices sélectionnées. Pour être efficace, il faudrait fixer un seuil supérieur à 0.5 de manière à obtenir une matrice à diagonale dominante, donc inversible. En pratique et compte tenu de l'application, un tel seuil n'est pas envisageable. En effet, selon la structure spatiale de la scène observée et le type de classes représenté, l'ensemble des pixels dont les proportions en label  $l$  est supérieure à un seuil donné peut être très restreint (voire vide selon la valeur du seuil). Par exemple, une classe telle que l'eau est généralement minoritaire dans un pixel basse résolution. Pour sélectionner des pixels capables de représenter toutes les classes, nous proposons de considérer un critère plus général fondé sur le contrôle du rapport défini pour tout label  $l \in \mathcal{L}$  et pour tout pixel  $y_i$  par

$$R_l(y_i) = \frac{\alpha_l(y_i)}{\max_{k \neq l} \alpha_k(y_i)}.$$

Pour chaque label, les pixels  $y_i$  pour lesquels le rapport  $R_l(y_i)$  est le plus élevé correspondent aux pixels qui représentent le label  $l$  en forte proportion ou tous les labels différents de  $l$  en faible proportion. Le choix d'un seuil étant contraignant et dépendant des données, nous proposons plutôt de considérer, pour chaque label, l'ensemble des  $n$  pixels  $y_i$  pour lesquels le rapport  $R_l(y_i)$  est le plus élevé. Cette approche permet de sélectionner les pixels dans un ensemble de taille  $n$ , où  $n \in [0, |\mathcal{D}_{BR}|]$ . Cette nouvelle approche est décrite par l'algorithme 6. La figure 8.7 montre

1. Pour  $l = 1$  à  $L$ 
  - Pour  $i = 1$  à  $|\mathcal{D}_{BR}|$ 
    - $\max = \max_{k \neq l} \alpha_k(y_i)$ ,
    - $R[l][y_i] = \frac{\alpha_l(y_i)}{\max}$ ,
  - fin pour  $i$ .
  - Trier le vecteur  $R[l]$  par ordre croissant,
  - $I[l] = (y_i)_{i=0 \dots |\mathcal{D}_{BR}|-1}$  tel que  $\forall i \leq j, R[l][y_i] \leq R[l][y_j]$ ,
2. fin pour  $l$ .
3. Pour  $l = 1$  à  $L$ 
  - tirer un entier  $m$  uniformément dans  $[|\mathcal{D}_{BR}| - n + 1, |\mathcal{D}_{BR}|]$ ,
  - le pixel correspondant  $y_l = I[l][m]$ ,
  - vérifier que ce pixel n'a pas été sélectionné deux fois.
4. fin pour  $l$ .

**Algorithme 6:** Sélection de sous-systèmes - stratégie 2.

l'efficacité de cette stratégie (figure (c)), pour  $n = 20$ , par rapport à l'algorithme 5. Les valeurs de conditionnement obtenues sont alors regroupées, pour la plupart, entre 0 et 3 (en échelle log) mais un pic d'occurrence persiste autour de 17 (*i.e.* l'infini). Ce pic montre qu'une proportion moindre mais non négligeable des matrices sélectionnées n'est pas inversible. Cependant, en diminuant la valeur du paramètre  $n$ , les performances peuvent être nettement améliorées. Par exemple, en se limitant à  $n = 15$ , les matrices sélectionnées sont toutes inversibles. La figure 8.8 présente les histogrammes obtenus avec cette stratégie pour trois différentes valeurs du paramètre  $n$  ( $n = 5, 10, 15$ ). Rappelons que le paramètre  $n$  indique le pourcentage des pixels de l'image susceptibles d'être sélectionnés pour l'estimation. Par exemple, pour une image de taille  $256 \times 256$ , fixer  $n = 5, 10, 15$  revient à choisir chaque pixel dans un échantillon comprenant, respectivement, 3276, 6553 ou 9830 pixels pour estimer 6 labels (dans le cas des expériences présentées). Avec  $n = 10$  (figure (b)), remarquons que les conditionnements des matrices sélectionnées sont essentiellement compris entre 0 et 1 (en échelle log), ce qui correspond aux trois premiers histogrammes présentés figure 8.6, où les occurrences de faibles valeurs de  $NFA$  sont les plus fréquentes. L'histogramme obtenu figure (a) pour  $n = 5$  est encore plus performant pour l'obtention de faibles valeurs de conditionnement, et donc du  $NFA$  (d'après la figure 8.6, 1er histogramme).

Pour traiter le cas des classes minoritaires, une autre approche serait de sélectionner un pixel parmi les plus représentatifs d'une classe  $l$  donnée et nettement plus représentatif du label  $l$

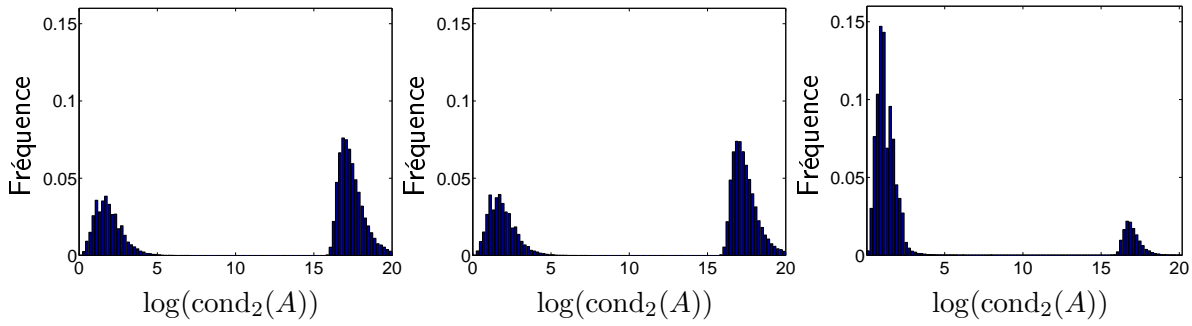


FIG. 8.7.: Histogramme du conditionnement (échelle log) pour la norme 2 obtenu pour 100 000 matrices  $A$  de taille  $6 \times 6$  correspondant réellement à des proportions et sélectionnées, figure (a), aléatoirement (uniformément), figure (b) avec la stratégie 1 (seuil=0.2) et, figure (c), avec la stratégie 2 pour  $n = 20$ .

que les autres pixels sélectionnés. Du point de vue de la matrice  $A$  des proportions, cette condition porte sur les colonnes (classes) et ne dicte pas directement une stratégie de sélection de pixels. Nous proposons, par conséquent, de chercher un critère portant sur les coefficients d'une matrice et capable de contrôler son conditionnement. Pour cela, nous introduisons la définition 8.4.1.

**Définition 8.4.1** Une matrice  $A$  de terme général  $(a_{ij})_{1 \leq i, j \leq n}$  est dite à diagonale  $\varepsilon$ -dominante en colonne pour la norme  $L^p$  si

$$\text{pour tout } j = 1, \dots, n, \quad \|A_j\|_p \leq \varepsilon |a_{jj}|,$$

avec  $A_j = {}^t(a_{1,j}, \dots, a_{j-1,j}, 0, a_{j+1,j}, \dots, a_{n,j})$ .

En considérant le conditionnement pour la norme  $L^1$ , le théorème 8.4.2 donne une majoration du conditionnement d'une matrice  $\varepsilon$ -dominante en colonne en fonction de ses coefficients diagonaux.

**Théorème 8.4.2** Soit  $A = (a_{ij})_{i,j}$  une matrice de taille  $n \times n$ . Si  $A$  est à diagonale  $\varepsilon$ -dominante en colonnes pour la norme  $L^1$ , avec  $\varepsilon < 1$ , alors le conditionnement de la matrice  $A$  pour la norme  $\|\cdot\|_1$  satisfait

$$\text{cond}_1(A) \leq \rho \cdot \frac{1 + \varepsilon}{1 - \varepsilon}, \quad (8.4.10)$$

$$\text{où } \rho = \frac{\max_{1 \leq i \leq n} |a_{ii}|}{\min_{1 \leq i \leq n} |a_{ii}|}.$$

**Démonstration.**

Soit  $D = (d_{ij})_{1 \leq i, j \leq n}$  la matrice diagonale où pour tout  $i$ ,  $d_{ii} = a_{ii}$ . La matrice  $A$  peut s'écrire sous la forme du produit  $AD^{-1}D$ .

**Lemme 8.4.3** Soient  $A$  et  $B$  deux matrices inversibles de taille  $n \times n$ , le conditionnement du produit  $AB$  pour la norme  $p$  vérifie l'inégalité

$$\text{cond}_p(AB) \leq \text{cond}_p(A) \text{cond}_p(B), \quad (8.4.11)$$

pour  $p = 1, +\infty$ .

**preuve du lemme** : Toute norme matricielle subordonnée à une norme vectorielle est sous-multiplicative, donc le conditionnement de la matrice  $AB$ , pour la norme subordonnée  $p$ , vérifie

$$\text{cond}_p(AB) = |||AB|||_p |||(AB)^{-1}|||_p \leq |||A|||_p |||B|||_p |||B^{-1}|||_p |||A^{-1}|||_p \quad (8.4.12)$$

$$\leq \text{cond}_p(A) \text{cond}_p(B). \quad (8.4.13)$$

D'après le lemme 8.4.3, le conditionnement de  $A$  peut être décomposé sous la forme :

$$\text{cond}_p(A) = \text{cond}_p(AD^{-1}D) \leq \text{cond}_p(AD^{-1}) \text{cond}_p(D). \quad (8.4.14)$$

La matrice  $D$  est diagonale donc, par définition,

$$\text{cond}_1(D) = \frac{\max_{1 \leq i \leq n} |a_{ii}|}{\min_{1 \leq i \leq n} |a_{ii}|}. \quad (8.4.15)$$

Ce rapport est noté  $\rho$ . On s'intéresse ensuite au conditionnement de la matrice  $AD^{-1}$ , notée  $C$ , et  $c_{ij}$  ses coefficients. Remarquons que, pour tout  $i$ ,  $c_{ii} = 1$  et pour tout  $j \neq i$ ,  $c_{ij} = \frac{a_{ij}}{a_{jj}}$ . Si  $A$  est  $\varepsilon$ -dominante en colonne pour la norme  $L^1$ , alors

$$\text{pour tout } j = 1, \dots, n, \quad \sum_{i=1, i \neq j}^n |a_{ij}| \leq \varepsilon |a_{jj}|.$$

Donc

$$\sum_{i=1, i \neq j}^n |c_{ij}| = \sum_{i=1, i \neq j}^n \frac{|a_{ij}|}{|a_{jj}|} \quad (8.4.16)$$

$$= \frac{1}{|a_{jj}|} \sum_{i=1, i \neq j}^n |a_{ij}| \quad (8.4.17)$$

$$\leq \varepsilon \frac{|a_{jj}|}{|a_{jj}|} = \varepsilon. \quad (8.4.18)$$

La norme  $|||\cdot|||_1$  de  $C$  vérifie alors

$$|||C|||_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |c_{ij}| \quad (8.4.19)$$

$$= \max_{1 \leq j \leq n} \left( \sum_{i=1, i \neq j}^n |c_{ij}| + |c_{ii}| \right) \quad (8.4.20)$$

$$\leq 1 + \varepsilon. \quad (8.4.21)$$

Quant à la norme de  $C^{-1}$ , utilisons le fait que  $|||C^{-1}|||_1 = |||({}^tC)^{-1}|||_\infty$ . Soit  $E = {}^tC$  la matrice de terme général  $e_{ij}$ . Par hypothèse, les coefficients de  $E$  vérifient alors l'inégalité :  $\sum_{j \neq i}^n e_{ij} < \varepsilon$  et, par conséquent,

$$|(Ex)_i| = \left| \sum_j e_{ij} x_j \right| \geq |e_{ii} x_i| - \left| \sum_{j \neq i}^n e_{ij} x_j \right| \quad (8.4.22)$$

$$\geq |x_i| - \left( \max_{j \neq i} |x_j| \right) \varepsilon. \quad (8.4.23)$$

La norme  $L^\infty$  du vecteur  $Ex$  satisfait alors la relation

$$\|Ex\|_\infty = \max_{1 \leq i \leq n} |(Ex)_i| \geq \|x\|_\infty - \varepsilon \|x\|_\infty \quad (8.4.24)$$

$$\geq (1 - \varepsilon) \|x\|_\infty, \quad (8.4.25)$$

donc

$$|||C^{-1}|||_1 = |||E^{-1}|||_\infty \leq \frac{1}{1 - \varepsilon}. \quad (8.4.26)$$

Ce dernier résultat permet de conclure que le conditionnement de la matrice  $A$  pour la norme  $|||\cdot|||_1$  vérifie

$$\text{cond}_1(A) \leq \rho \cdot \frac{1 + \varepsilon}{1 - \varepsilon}. \quad (8.4.27)$$

□

Une stratégie de sélection de matrices fondée sur une condition de  $\varepsilon$ -dominance permettrait alors de contrôler leur conditionnement. Plus précisément, dans l'objectif de sélectionner des matrices  $\varepsilon$ -dominantes en colonne, nous proposons de sélectionner les pixels  $y_i$  de manière à vérifier les conditions, pour  $n$  fixé :

$$\begin{aligned} & \cdot \alpha_i(y_i) \geq P_i^{1-n}, \\ & \cdot \forall k \neq i, \alpha_k(y_i) \leq P_k^n, \end{aligned}$$

où  $P_k^n$  correspond au  $n$ ème percentile de  $\{\alpha_k(y_i)\}_i$ . L'algorithme 7 permet sélectionner des matrices qui vérifient ces conditions. Le paramètre  $n$  joue le même rôle que dans l'algorithme 6.

La figure 8.9 présente les histogrammes obtenus pour  $n = 5, 10, 15$ . Si ces histogrammes sont comparables à ceux de la figure 8.8, les valeurs de conditionnement obtenues ont tendance à être plus élevées. Cette stratégie est donc plutôt moins performante pour ce jeu de données, même si les résultats restent globalement assez proches. Cependant, une étude plus poussée serait nécessaire pour étudier les performances de ces deux algorithmes sur des données très différentes. En effet, nous présentons ici une étude restreinte à un seul exemple d'images alors que l'objectif des algorithmes de sélection que nous proposons est de permettre le traitement d'images comprenant une ou plusieurs classes minoritaires. Si les résultats présentés ne sont donc pas suffisants pour conclure sur les performances relatives de ces deux stratégies, ils permettent d'ores et déjà de réduire nettement le nombre d'itérations et, par conséquent, le temps de calcul nécessaire à la convergence de l'algorithme.

En pratique, dans les expériences qui suivent, nous utilisons la stratégie 2 (Algorithme 6) pour sélectionner les sous-systèmes à inverser pour l'estimation des moyennes des classes. L'inversion est ensuite réalisée par décomposition LU. Dans le chapitre 9, nous présentons les performances théoriques du modèle de détection en fonction du contraste et de la taille de l'image, ainsi que du pourcentage de changements présents.

1. Pour  $l = 1$  à  $L$ 
  - trier le vecteur des proportions  $(\alpha_l(y_i))_{i=1\dots|\mathcal{D}_{BR}|}$  par ordre croissant,
  - $(I_l(i))_{i=1\dots|\mathcal{D}_{BR}|}$  : indice du pixel en  $i$ ème position pour le label  $l$  dans le vecteur des proportions,
  - $(I_l^{-1}(i))_{i=1\dots|\mathcal{D}_{BR}|}$  : rang du pixel d'indice  $i$  dans le vecteur des proportions ordonné pour le label  $l$ .
2. fin pour.
3. Pour  $l = 1$  à  $L$ 
  - calculer  $\forall i, p_l(i) = |\mathcal{D}_{BR}| + 1 - I_l^{-1}(i)$ ,
  - calculer  $\forall i, q_l(i) = \max_{l' \neq l} I_{l'}^{-1}(i)$ ,
  - calculer  $\forall i, \mathcal{P}_l(i) = \max(p_l(i), q_l(i))$ ,
  - trier le vecteur  $(\mathcal{P}_l(i))_{i=1\dots|\mathcal{D}_{BR}|}$  par ordre croissant,
  - $(J_l(i))_{i=1\dots|\mathcal{D}_{BR}|}$  : indice du pixel en  $i$ ème position dans le vecteur  $(\mathcal{P}_l(i))_i$ ,
4. fin pour.
5. Pour  $l = 1$  à  $L$ 
  - tirer un entier  $k$  uniformément dans  $[1, m]$ ,
  - le pixel correspondant  $y_l = J_l(k)$ ,
  - vérifier que le pixel n'a pas été sélectionné deux fois.
6. fin pour.

**Algorithme 7:** Sélection de sous-systèmes - stratégie 3.

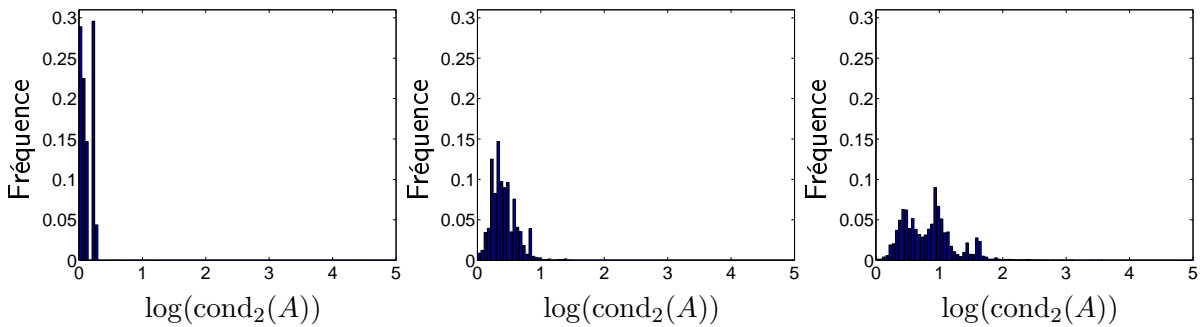


FIG. 8.8.: Histogramme du conditionnement (échelle log) pour la norme 2 obtenu pour 100 000 matrices  $A$  de taille  $6 \times 6$  correspondant réellement à des proportions et sélectionnées selon la stratégie 2, avec  $n = 5$  figure (a), avec  $n = 10$  figure (b) et avec  $n = 15$  figure (c).

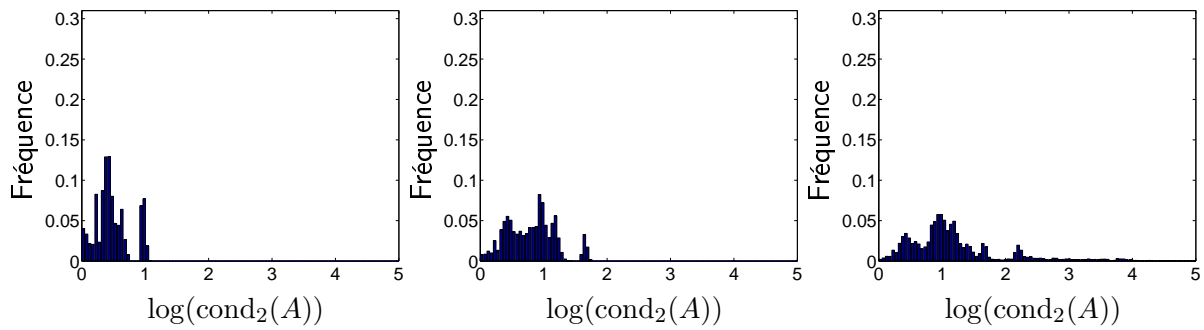


FIG. 8.9.: Histogramme du conditionnement (échelle log) pour la norme 2 obtenu pour 100 000 matrices  $A$  de taille  $6 \times 6$  correspondant réellement à des proportions et sélectionnées selon la stratégie 3, avec  $n = 5$  figure (a), avec  $n = 10$  figure (b) et avec  $n = 15$  figure (c).

## 9. Performances théoriques

Dans ce chapitre, nous nous intéressons au comportement du modèle en fonction des différents paramètres intrinsèques aux images traitées (contraste, taille de l'image, facteur de résolution, etc.). Pour des raisons de simplicité, cette étude est réalisée dans le contexte monodimensionnel.

### 9.1. Modélisation des données

Dans cette section, nous décrivons un modèle simple d'images permettant de séparer les paramètres importants pour la détection avant d'étudier les performances du modèle décrit dans la section 7 vis-à-vis de chacun de ces paramètres. Pour commencer, nous considérons une image  $v$  de même résolution spatiale que la classification. Supposons que toute image  $u$  puisse s'écrire sous la forme

$$u = I + b,$$

où  $I$  est une image constante par morceaux et  $b$  est une image de bruit gaussien. Etant donnée une classification contenant  $L$  labels, supposons que l'image  $I$  est parfaitement superposable à l'image des labels, avec en chaque pixel la valeur caractéristique du label correspondant. La variance empirique  $\sigma^2$  de l'image  $u$  s'obtient directement par la relation  $\sigma^2 = \sigma_I^2 + \sigma_b^2$ . Pour analyser la sensibilité du nombre de fausses alarmes, l'estimation de la famille  $\mu$  des moyennes caractéristiques des classes est supposée exacte. L'erreur quadratique mesurée en moyenne pour chaque pixel vaut alors  $\sigma_b^2$ , les erreurs quadratiques cumulées sur un sous-domaine  $D$  valent en moyenne,

$$\mathbb{E} [\delta_D^2] = |D| \times \sigma_b^2.$$

D'après le corollaire 7.3.4, le *nombre de fausses alarmes* associé à une image  $v$  et une classification contenant  $L$  labels est obtenu, pour tout sous-domaine  $D$  de cardinal supérieur à  $|\mathcal{L}|$ , pour tout résidu  $\delta(u_D)$  et écart-type  $\sigma$  par

$$NFA(|D|, \delta(u_D), \sigma) = |\mathcal{D}_{BR}| C_{|\mathcal{D}_{BR}|}^{|D|} \Gamma_{inc}\left(\frac{|D| - |\mathcal{L}|}{2}, \frac{\delta(u_D)^2}{2\sigma^2}\right). \quad (9.1.1)$$

Pour alléger les notations, notons  $k = |D|$ ,  $n = |\mathcal{D}_{BR}|$  et considérons l'erreur moyenne  $\delta_k$  obtenue sur un domaine de taille  $k$  :  $\delta_k^2 = k \sigma_b^2$ . Le *nombre de fausses alarmes* d'un sous-domaine de



cardinal  $k$  contenu dans un domaine de cardinal  $n$  s'exprime alors, pour  $\delta(u_D) = \delta_k$ , par

$$\begin{aligned}
 NFA(k, \delta_k, \sigma) &= nC_n^k \Gamma_{inc} \left( \frac{k-L}{2}, \frac{\delta_k^2}{2\sigma^2} \right) \\
 &= nC_n^k \Gamma_{inc} \left( \frac{k-L}{2}, \frac{k\sigma_b^2}{2(\sigma_I^2 + \sigma_b^2)} \right) \\
 &= nC_n^k \Gamma_{inc} \left( \frac{k-L}{2}, \frac{k}{2(\frac{\sigma_I^2}{\sigma_b^2} + 1)} \right). \tag{9.1.2}
 \end{aligned}$$

Cette simplification permet de remarquer que le nombre de fausses alarmes dépend essentiellement de la taille du sous-domaine considéré, du nombre de labels présents dans la labellisation et du rapport de variances  $\frac{\sigma_I}{\sigma_b}$ . Ce rapport peut être vu comme une mesure de contraste de l'image considérée. Notons  $c = \frac{\sigma_I}{\sigma_b}$  le contraste de l'image  $v$  ( $c \geq 0$ ), et introduisons le paramètre  $\gamma = \frac{1}{c^2+1}$ ,  $\gamma \in ]0, 1]$ . D'autre part, le cardinal  $k$  du sous-domaine étudié peut être vu comme un certain pourcentage  $p \in [0, 1]$  de l'ensemble des  $n$  pixels de l'image, i.e.  $p = k/n$ . Le nombre de fausses alarmes d'une image de contraste  $c$  sur un sous-domaine représentant une proportion  $p$  du domaine étudié (de cardinal  $n$ ) peut alors s'exprimer en fonction de ces nouveaux paramètres sous la forme d'une fonction  $\mathcal{F}$  définie par

$$\begin{aligned}
 \mathcal{F}(p, n, \gamma) &= nC_n^{pn} \Gamma_{inc} \left( \frac{pn-L}{2}, \gamma \frac{pn}{2} \right) \\
 &= nC_n^{pn} \frac{1}{\Gamma(\frac{pn-L}{2})} \int_0^{\gamma \frac{pn}{2}} e^{-t} t^{\frac{pn-L}{2}-1} dt, \tag{9.1.3}
 \end{aligned}$$

pour tout  $n \in \mathbb{N}$ , pour tout  $p \in ]\frac{L}{n}, 1]$ , et  $\gamma = \frac{1}{c^2+1} \in ]0, 1]$ . Cette nouvelle expression nous permet d'étudier la significativité d'un sous-domaine basse résolution en fonction du contraste de l'image considérée, de sa taille et du pourcentage de pixels de changements. En particulier, cette étude permet d'accéder aux valeurs limites de ces paramètres pour lesquels le nombre de fausses alarmes est encore significatif.

La section suivante est centrée sur la sensibilité du  $NFA$  au contraste de l'image et, en particulier, sur la détectabilité d'un domaine en fonction de son niveau de contraste.

## 9.2. Sensibilité au niveau de contraste

La comparaison d'images est, en général, nettement améliorée lorsque les images ont la même dynamique de niveaux de gris, d'autant plus que la plupart des méthodes de comparaison sont basées sur l'hypothèse que deux points en correspondance spatiale ont la même intensité. Les valeurs acquises par satellite sont très variables selon les conditions d'illumination, les mesures peuvent donc varier fortement en fonction de la date (et heure) d'acquisition, et en fonction des saisons. Typiquement, les conditions d'ensoleillement hivernales donnent lieu, en général, à des images au contraste faible. Une méthode de détection de changements doit être performante

quelque soit la date d'acquisition et, par conséquent, quelque soit le contraste de l'image, en particulier lorsque le contraste est faible.

Tout d'abord, intéressons-nous au cas particulier où le contraste de l'image est nul ( $c = 0$ ). Ce cas correspond, pour notre modèle d'image, à un rapport d'écart-type  $\frac{\sigma_I}{\sigma_b} = 0$ . Un tel cas de figure peut se produire si l'image est tellement bruitée que sa structure géométrique est *noyée* dans le bruit (*i.e.*  $\sigma_I$  est quelconque et  $\sigma_b = +\infty$ ), ou alors et quelque soit le niveau de bruit, si l'image n'a pas de structure géométrique ( $\sigma_I = 0$ ). Le modèle *a contrario* a été choisi de manière à garantir l'absence de toute détection dans ce type d'images (*c.f* Chapitre 7, section 7.3). Cette propriété est vérifiée analytiquement par la proposition 9.2.1, car lorsque le contraste de l'image est nul, le paramètre  $\gamma$  vaut 1.

**Proposition 9.2.1** *Si  $\gamma = 1$ , pour  $n > 1$  fixés et pour tout  $p \in ]\frac{L}{n}, 1]$ ,  $\mathcal{F}(p, n, 1) \geq 1$ .*

**Démonstration.**

Si  $\gamma = 1$ ,

$$\mathcal{F}(p, n, 1) = nC_n^{pn} \frac{1}{\Gamma(\frac{pn-L}{2})} \int_0^{\frac{pn}{2}} e^{-t} t^{\frac{pn-L}{2}-1} dt. \quad (9.2.4)$$

Commençons par calculer la partie intégrale de (9.2.4). Pour plus de lisibilité, posons  $x = \frac{pn}{2}$  et  $a = \frac{L}{2}$ . Pour tout  $a > 0$  et pour tout  $x \geq a$ , une intégration par parties multiple permet d'écrire

$$\begin{aligned} \int_0^x e^{-t} t^{x-a-1} dt &= \left[ - \sum_{q=1}^{x-a-1} (x-a-1) \cdots (x-a-q+1) t^{x-a-q} e^{-t} \right]_0^x \\ &\quad + (x-a-1)! \int_0^x e^{-t} dt \\ &= - \sum_{q=1}^{x-a-1} \frac{(x-a-1)!}{(x-a-q)!} x^{x-a-q} e^{-x} + (x-a-1)! (1 - e^{-x}) \\ &= (x-a-1)! \left( 1 - \sum_{q=1}^{x-a} \frac{x^{x-a-q}}{(x-a-q)!} e^{-x} \right) \\ &= (x-a-1)! \left( 1 - e^{-x} \sum_{q=0}^{x-a-1} \frac{x^q}{q!} \right) \\ &= (x-a-1)! \left( e^{-x} \sum_{q=x-a}^{+\infty} \frac{x^q}{q!} \right). \end{aligned} \quad (9.2.5)$$

Par ailleurs, remarquons que, pour tout  $x \in \mathbb{N}$  et  $k \in \mathbb{N}$  tel que  $x-1-k \geq 0$ ,

$$\frac{x^{x-1-k}}{(x-1-k)!} \leq \frac{x^{x+k}}{(x+k)!}. \quad (9.2.6)$$

En effet,

$$\begin{aligned}
 \frac{(x+k)!}{(x-1-k)!} &= (x+k)(x+k-1)\cdots(x-k) \\
 &= (x^2-k^2)(x^2-(k-1)^2)\cdots(x^2-1)x \\
 &\leq x^{2k+1} = \frac{x^{x+k}}{x^{x-1-k}}.
 \end{aligned} \tag{9.2.7}$$

Donc

$$\sum_{k=0}^{x-1} \frac{x^{x-1-k}}{(x-1-k)!} \leq \sum_{k=0}^{x-1} \frac{x^{x+k}}{(x+k)!}. \tag{9.2.8}$$

Le changement d'indice  $q = x + k$  permet d'écrire

$$\sum_{k=0}^{x-1} \frac{x^{x+k}}{(x+k)!} = \sum_{q=x}^{2x-1} \frac{x^q}{q!} \leq \sum_{q=x}^{+\infty} \frac{x^q}{q!},$$

et, avec le changement d'indice  $q = x - 1 - k$ ,

$$\sum_{k=0}^{x-1} \frac{x^{x-1-k}}{(x-1-k)!} = \sum_{q=0}^{x-1} \frac{x^q}{q!}.$$

La relation (9.2.8) implique alors

$$\sum_{q=0}^{x-1} \frac{x^q}{q!} \leq \sum_{q=x}^{+\infty} \frac{x^q}{q!}.$$

De plus, en remarquant que

$$\sum_{q=0}^{x-1} \frac{x^q}{q!} + \sum_{q=x}^{+\infty} \frac{x^q}{q!} = e^x,$$

nous obtenons la majoration

$$\sum_{q=x}^{+\infty} \frac{x^q}{q!} \geq \frac{1}{2}e^x, \tag{9.2.9}$$

et donc

$$e^{-x} \sum_{q=x}^{+\infty} \frac{x^q}{q!} \geq \frac{1}{2}. \tag{9.2.10}$$

L'intégrale (9.2.5) satisfait alors

$$\begin{aligned}
 \int_0^x e^{-t} t^{x-a-1} dt &\geq \Gamma(x-a) e^{-x} \sum_{q=x}^{+\infty} \frac{x^q}{q!} \\
 &\geq \frac{\Gamma(x-a)}{2}.
 \end{aligned} \tag{9.2.11}$$

Cette minoration permet de conclure

$$\begin{aligned}
 \mathcal{F}(p, n, 1) &= n C_n^{pn} \frac{1}{\Gamma(\frac{pn-L}{2})} \int_0^{\frac{pn}{2}} e^{-t} t^{\frac{pn-L}{2}-1} dt \\
 &\geq \frac{n}{2} C_n^{pn},
 \end{aligned} \tag{9.2.12}$$

i.e. dès que  $n > 1$ , et  $p > \frac{L}{n}$ ,  $\mathcal{F}(p, n, 1) \geq 1$ .  $\square$

Plus généralement, nous nous intéressons à l'évolution du *nombre de fausses alarmes* lorsque le contraste de l'image varie (9.2.2) et atteint des valeurs limites (9.2.3).

**Proposition 9.2.2** *Pour tout  $n \in \mathbb{N}^*$  et  $p \in ]\frac{L}{n}, 1]$  fixés, la fonction  $\gamma \mapsto \mathcal{F}(p, n, \gamma)$  est croissante.*

**Démonstration.**

Soient  $n > 0$  et  $p \in ]\frac{L}{n}, 1]$  fixés, la dérivée de la fonction  $\mathcal{F}$  par rapport au paramètre  $\gamma$

$$\begin{aligned} \frac{\partial \mathcal{F}}{\partial \gamma}(p, n, \gamma) &= \frac{n^2 p C_n^{pn}}{2\Gamma(\frac{pn-L}{2})} e^{-\gamma \frac{pn}{2}} \left(\frac{\gamma pn}{2}\right)^{\frac{pn-L}{2}-1} \\ &\geq 0. \end{aligned}$$

La fonction  $\gamma \mapsto \mathcal{F}(p, n, \gamma)$  est donc croissante.  $\square$

Le paramètre  $\gamma = \frac{1}{c^2+1}$  est inversement proportionnel au contraste de l'image. Donc ce résultat peut être interprété, du point de vue du *NFA*, par le fait qu'un sous-domaine quelconque représentant une proportion  $p$  donnée du domaine basse résolution est d'autant plus significatif que l'image est contrastée. La cohérence mesurée entre une classification et une image  $\mathcal{D}_{BR}$  est d'autant plus forte que son contraste est élevé, ce qui signifie qu'un sous-domaine est d'autant mieux validé par le *NFA* que son contraste est fort.

A l'inverse, étudions maintenant le comportement du *NFA* lorsque le bruit est très faible dans l'image ou que sa structure géométrique est très forte devant le bruit. Le contraste  $c$  de l'image tend alors vers l'infini, et le paramètre  $\gamma$  vers 0.

**Proposition 9.2.3** *Pour tout  $n > 0$  et  $p \in ]\frac{L}{n}, 1]$  fixés,  $\lim_{\gamma \rightarrow 0} \mathcal{F}(p, n, \gamma) = 0$ .*

**Démonstration.**

L'application  $\gamma \mapsto \mathcal{F}(p, n, \gamma)$  est  $\mathcal{C}^\infty$  donc  $\lim_{\gamma \rightarrow 0} \mathcal{F}(p, n, \gamma) = \mathcal{F}(p, n, 0) = 0$ .  $\square$

Du point de vue du *NFA*, cette propriété indique que tout domaine est détectable dès que le contraste de l'image est suffisamment fort. Bien qu'intuitif, ce résultat prouve en particulier que des images contenant des changements en forte proportion peuvent être analysées si leur contraste est suffisamment fort. Dans la section suivante, nous présentons une analyse asymptotique du *NFA* en fonction de la taille des images.

### 9.3. Influence de la taille des images

Après avoir analysé la sensibilité du *NFA* vis à vis des variations de contraste, nous proposons d'étudier l'existence de sous-domaines détectables pour un niveau de contraste fixé, ainsi que l'amélioration de la détectabilité lorsque la taille de l'image grandit.

**Proposition 9.3.1** Pour tout  $\gamma$  fixé dans l'intervalle  $]0, 1[$ , il existe  $p^*(\gamma)$  tel que, pour tout  $p \in [p^*(\gamma), 1[$ ,

$$\lim_{n \rightarrow +\infty} \mathcal{F}(p, n, \gamma) = 0.$$

**Démonstration.**

Rappelons que, pour tout  $n > 1$ ,  $p \in ]\frac{L}{n}, 1]$  et  $\gamma \in ]0, 1[$ ,

$$\mathcal{F}(p, n, \gamma) = n C_n^{pn} \frac{1}{\Gamma(\frac{pn-L}{2})} \int_0^{\gamma \frac{pn}{2}} e^{-t} t^{\frac{pn-L}{2}-1} dt \quad (9.3.13)$$

Posons  $a = \frac{pn-L}{2} - 1$ . La formule de Stirling donne l'équivalence

$$\Gamma(a+1) \underset{a \rightarrow \infty}{\sim} \sqrt{2\pi a} \left(\frac{a}{e}\right)^a. \quad (9.3.14)$$

La fonction  $\mathcal{F}$  s'écrit alors

$$\begin{aligned} \mathcal{F}(p, n, \gamma) &= C_n^{pn} \frac{2(a+1+\frac{L}{2})}{p\Gamma(a+1)} \int_0^{\gamma(a+1+\frac{L}{2})} e^{-t} t^a dt \\ &\underset{a \rightarrow \infty}{\sim} C_n^{pn} \int_0^{\gamma(a+1+\frac{L}{2})} e^{-t} t^a \frac{2(a+1+\frac{L}{2})}{p\sqrt{2\pi a}} \left(\frac{e}{a}\right)^a dt. \end{aligned} \quad (9.3.15)$$

Avec le changement de variable  $u = \frac{t}{a}$ , on obtient

$$\begin{aligned} \mathcal{F}(p, n, \gamma) &= C_n^{pn} \int_0^{\gamma(1+\frac{1}{a}+\frac{L}{2a})} e^{-au} (au)^a \frac{2(a+1+\frac{L}{2})}{p\sqrt{2\pi a}} \left(\frac{e}{a}\right)^a a du \\ &= C_n^{pn} \frac{2a(a+1+\frac{L}{2})}{p\sqrt{2\pi a}} \int_0^{\gamma(1+\frac{1}{a}+\frac{L}{2a})} (e^{1-u} u)^a du. \end{aligned} \quad (9.3.16)$$

La fonction  $u \mapsto e^{1-u} u$  admet pour dérivée la fonction  $u \mapsto e^{1-u}(1-u)$ . Elle est donc strictement croissante sur l'intervalle  $[0, 1[$ . De plus, il existe un rang  $A$  à partir duquel  $\forall a > A$ ,  $\frac{a}{a+1+\frac{L}{2}} > \gamma$ ,

i.e.  $\gamma \frac{a+1+\frac{L}{2}}{a} < 1$ . Par conséquent, la fonction  $\mathcal{F}$  vérifie la majoration

$$\begin{aligned} \mathcal{F}(p, n, \gamma) &\leq C_n^{pn} \frac{2a(a+1+\frac{L}{2})}{p\sqrt{2\pi a}} \frac{\gamma(a+1+\frac{L}{2})}{a} \left( e^{1-\gamma \frac{a+1+\frac{L}{2}}{a}} \gamma \frac{a+1+\frac{L}{2}}{a} \right)^a \\ &\leq C_n^{pn} \frac{2\gamma(a+1+\frac{L}{2})^2}{p\sqrt{2\pi a}} e^{a-\gamma(a+1+\frac{L}{2})+a \ln(\gamma \frac{a+1+\frac{L}{2}}{a})} \\ &\leq C_n^{pn} \frac{2\gamma(a+1+\frac{L}{2})^2}{p\sqrt{2\pi a}} e^{a(1-\gamma+\ln \gamma+\ln(1+\frac{1}{a}+\frac{L}{2a}))-\gamma-\frac{\gamma L}{2}} \\ &\leq C_n^{pn} \frac{2\gamma(a+1+\frac{L}{2})^2}{p\sqrt{2\pi a}} e^{a(1-\gamma+\ln \gamma)+(1-\gamma)(1+\frac{L}{2})+o(1)} \end{aligned} \quad (9.3.17)$$

car  $\ln(1+\frac{1}{a}+\frac{L}{2a}) = \frac{1}{a} + \frac{L}{2a} + o(\frac{1}{a})$  au voisinage de  $+\infty$ .

Si  $p = 1$ , alors

$$\mathcal{F}(1, n, \gamma) \leq \frac{2\gamma(a+1+\frac{L}{2})^2}{\sqrt{2\pi a}} e^{a(1-\gamma+\ln \gamma)+(1-\gamma)(1+\frac{L}{2})+o(1)}. \quad (9.3.18)$$

Considérons, pour tout  $\gamma \in ]0, 1[$ , la fonction  $\phi : \gamma \mapsto 1 - \gamma + \ln \gamma$ . Elle admet pour dérivée la fonction  $\gamma \mapsto -1 + \frac{1}{\gamma}$ , positive sur  $]0, 1[$ . De plus, la fonction  $\phi$  tend vers  $-\infty$  lorsque  $\gamma$  tend vers 0, et  $\phi(1) = 0$ . La fonction  $\phi$  est donc croissante et strictement négative sur l'intervalle  $]0, 1[$ . Par conséquent,

$$\lim_{a \rightarrow \infty} \frac{2\gamma(a+1)^2}{\sqrt{2\pi a}} e^{a(1-\gamma+\ln \gamma)+(1-\gamma)(1+\frac{L}{2})+o(1)} = 0. \quad (9.3.19)$$

Or  $a = \frac{pn-L}{2} - 1$  et donc  $\lim_{n \rightarrow \infty} \mathcal{F}(p, n, \gamma) = 0$ . Si  $p \neq 1$ , utilisons les propriétés du logarithme pour écrire

$$\ln(C_n^{pn}) = \ln(n!) - \ln((pn)!) - \ln((n-pn)!),$$

et

$$\ln(n!) = n(\ln n - 1) + o(n),$$

d'où

$$\begin{aligned} \ln(C_n^{pn}) &= n(\ln n - 1) + o(n) - pn(\ln(pn) - 1) + o(pn) \\ &\quad - (n-pn)(\ln(n-pn) - 1) + o(n-pn) \end{aligned} \quad (9.3.20)$$

$$\begin{aligned} &= n \ln n - n - pn \ln p - pn \ln n + pn - n \ln n - n \ln(1-p) \\ &\quad + pn \ln n + pn \ln(1-p) + n - pn + o(n) \\ &= -pn \ln p - n(1-p) \ln(1-p) + o(n). \end{aligned} \quad (9.3.21)$$

Et, en remplaçant  $a$  par sa valeur  $\frac{pn-L}{2} - 1$  dans l'équation (9.3.17), alors

$$\begin{aligned} \mathcal{F}(p, n, \gamma) &\leq C_n^{pn} \frac{2\gamma p^2 n^2}{4p\sqrt{\pi(pn-L-2)}} e^{(\frac{pn-L}{2}-1)(1-\gamma+\ln \gamma)+(1-\gamma)(1+\frac{L}{2})+o(1)} \\ &\leq \frac{\gamma p n^2}{2\sqrt{\pi(pn-L-2)}} e^{-pn \ln p - n(1-p) \ln(1-p) + o(n)} e^{(\frac{pn-L}{2}-1)(1-\gamma+\ln \gamma)+(1-\gamma)(1+\frac{L}{2})+o(1)} \\ &\leq \frac{\gamma p n^2}{2\sqrt{\pi(pn-L-2)}} e^{n(p(\frac{1}{2}(1-\gamma+\ln \gamma) - \ln p) - (1-p) \ln(1-p) + o(1)) - \frac{L+2}{2}(1-\gamma+\ln \gamma)+(1-\gamma)(1+\frac{L}{2})+o(1)} \\ &\leq \frac{\gamma p n^2}{2\sqrt{\pi(pn-L-2)}} e^{n(p(\frac{1}{2}(1-\gamma+\ln \gamma) - \ln p) - (1-p) \ln(1-p) + o(1)) - \frac{L+2}{2} \ln \gamma + o(1)}. \end{aligned} \quad (9.3.22)$$

Posons

$$g(\gamma, p) = p \left( \frac{1}{2}(1 - \gamma + \ln \gamma) - \ln p \right) - (1-p) \ln(1-p),$$

et étudions les domaines de  $\gamma$  et  $p$  tels que  $g(\gamma, p) \leq 0$ .

La dérivée de la fonction  $g$  selon le paramètre  $p$ , pour  $p \neq 0$ , est définie par

$$\frac{\partial g(\gamma, p)}{\partial p} = \frac{1}{2}(1 - \gamma + \ln \gamma) + \ln \frac{1-p}{p}. \quad (9.3.23)$$

Remarquons que

$$\begin{aligned} \frac{\partial g(\gamma, p)}{\partial p} = 0 &\Leftrightarrow -\frac{1}{2}(1 - \gamma + \ln \gamma) = \frac{1-p}{p} \\ &\Leftrightarrow p(1 + e^{-\frac{1}{2}(1-\gamma+\ln \gamma)}) = 1 \\ &\Leftrightarrow p = \frac{1}{1 + e^{-\frac{1}{2}(1-\gamma+\ln \gamma)}}. \end{aligned} \quad (9.3.24)$$

Si  $p \leq \frac{1}{1+e^{-\frac{1}{2}(1-\gamma+\ln \gamma)}}$  alors la dérivée  $\frac{\partial g(\gamma, p)}{\partial p}$  est positive. De plus, pour tout  $\gamma$  fixé dans l'intervalle  $]0, 1[$ , la limite de la fonction  $g$  lorsque  $p$  tend vers 0 est nulle. Donc la fonction  $g(\gamma, p)$  est croissante selon  $p$  et positive quelque soit  $p$  dans l'intervalle  $[0, \frac{1}{1+e^{-\frac{1}{2}(1-\gamma+\ln \gamma)}}]$ . Pour tout  $\gamma$  fixé dans l'intervalle  $]0, 1[$ ,  $\lim_{p \rightarrow 1} g(\gamma, p) = \frac{1}{2}(1 - \gamma + \ln \gamma) < 0$  donc la fonction  $g$ , décroissante sur l'intervalle  $[\frac{1}{1+e^{-\frac{1}{2}(1-\gamma+\ln \gamma)}}, 1]$ , passe par 0 sur cet intervalle. Par conséquent, pour tout  $\gamma$  fixé dans  $]0, 1[$ , il existe une valeur  $p^*(\gamma)$  telle que, quelque soit  $p \in [p^*(\gamma), 1]$ ,  $g(\gamma, p) < 0$  et alors la limite du majorant lorsque  $n$  tend vers l'infini vaut

$$\lim_{n \rightarrow +\infty} \frac{\gamma p n^2}{2\sqrt{\pi(pn - L - 2)}} e^{n(p(\frac{1}{2}(1-\gamma+\ln \gamma) - \ln p) - (1-p)\ln(1-p) + o(1)) - \frac{L+2}{2}\ln \gamma + o(1)} = 0. \quad (9.3.25)$$

Ce résultat permet de conclure que, pour tout  $\gamma$  fixé dans l'intervalle  $]0, 1[$ , il existe une valeur  $p^*(\gamma)$  telle que, pour tout  $p$  appartenant à l'intervalle  $[p^*(\gamma), 1]$ ,  $\lim_{n \rightarrow +\infty} \mathcal{F}(p, n, \gamma) = 0$ . □

Pour un pourcentage  $p$  fixé de pixels cohérents avec la labellisation, la fonction  $g$  est croissante en fonction de  $\gamma$ . La figure 9.1 trace les valeurs de  $g$  lorsque le pourcentage  $p$  varie, pour différents niveaux du paramètre  $\gamma$  dépendants du contraste,

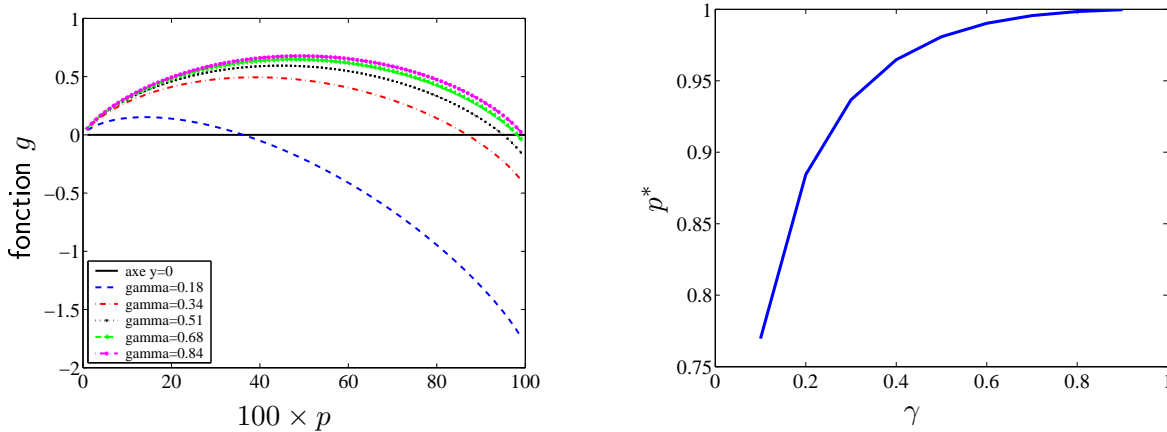


FIG. 9.1.: A gauche, chaque courbe représente la fonction  $g$  en fonction de  $p$  pour différentes valeurs de  $\gamma$ . A droite, la valeur  $p^*$  obtenue numériquement est tracée en fonction du paramètre  $\gamma$ .

Ce résultat peut être interprété par le fait que, pour un niveau de contraste donné, *i.e.*  $\gamma$  fixé, la méthode *a contrario* permet de détecter tous les changements présents dans une image s'ils représentent jusqu'à  $1 - p^*(\gamma)$  % de l'image, du moment que l'image est suffisamment grande.

Avec le modèle d'image utilisé pour cette analyse (section 9.1), le facteur de résolution n'apparaît pas directement comme un paramètre du modèle *a contrario*. Il intervient en fait essentiellement à travers la variance  $\sigma^2$  de l'image **BR**. Dans la section 9.4, nous tentons de définir la variance d'une image en fonction de sa résolution de manière à faire apparaître le facteur de résolution comme un paramètre du *nombre de fausses alarmes*.

## 9.4. Etude en fonction du facteur de résolution

En vue d'étudier plus précisément l'évolution du *nombre de fausses alarmes* en fonction du facteur de résolution considéré, nous cherchons à définir la variance d'une image **BR** en fonction de sa résolution.

Le facteur de résolution entre deux images  $u$  et  $v$  peut être vu comme un facteur de dégradation spatiale, ou encore en tant que taille de fenêtre dans lesquels l'image  $u$  est moyennée pour obtenir  $v$ . Ainsi, plus le facteur de résolution est fort, plus l'image  $v$  apparaît lisse comparativement à l'image  $u$ , *i.e.* plus la variance de  $v$  est faible devant celle de  $u$ . D'un point de vue plus concret, considérons une image  $u$  représentant une scène agricole, à une résolution donnée, suffisamment bonne pour permettre de distinguer les parcelles. Soit une image  $v$  de la même scène, plus le facteur de résolution entre  $u$  et  $v$  est fort, plus les pixels de l'image  $v$  sont mixtes, *i.e.* plus les labels représentés dans un pixel sont susceptibles d'être nombreux. L'évolution de la variance en fonction du facteur de dégradation vient donc non seulement du moyennage du bruit mais aussi du mélange des motifs présents dans l'image. Reprenons le modèle d'image adopté section 9.1 selon lequel une image quelconque  $u$  peut être considérée comme la somme de deux images indépendantes, l'une contenant l'information géométrique ( $I$ ), constante par morceaux, et l'autre le bruit ( $b$ ) supposé gaussien. Notons

$$u = I + b.$$

Soient  $N = \frac{|\mathcal{D}_{HR}|}{|\mathcal{D}_{BR}|}$  le rapport de résolution considéré,  $u(N)$ ,  $I(N)$  et  $b(N)$  respectivement les images  $u$ ,  $I$  et  $b$  dégradées d'un facteur  $N$  et  $\sigma_{u(N)}^2$ ,  $\sigma_{I(N)}^2$ ,  $\sigma_{b(N)}^2$ ,  $\sigma^2$ ,  $\sigma_{(I)}^2$  et  $\sigma_{(b)}^2$  les variances respectives des images  $u(N)$ ,  $I(N)$ ,  $b(N)$ ,  $u$ ,  $I$  et  $b$ . Les images  $I$  et  $b$  étant supposées indépendantes, les images  $I(N)$  et  $b(N)$  le sont aussi et la variance de  $u(N)$  est égale à la somme des variances de  $I(N)$  et de  $b(N)$ , donc

$$\sigma_{u(N)}^2 = \sigma_{I(N)}^2 + \frac{\sigma_b^2}{N}. \quad (9.4.26)$$

Du point de vue de l'application, l'image  $I$  est superposable à la labellisation avec, par exemple, la valeur en chaque pixel égale à la moyenne du label correspondant. Même si ce n'est pas réaliste, supposons dans un premier temps que chaque label est équidistribué. La variance de l'image  $I$  dégradée vaut alors

$$\sigma_{I(N)}^2 = \frac{\sigma_I^2}{N}, \quad (9.4.27)$$

et, donc, la variance de l'image  $u$  dégradée vaut simplement

$$\sigma_{u(N)}^2 = \frac{\sigma^2}{N}. \quad (9.4.28)$$

Comme dans la section 9.1, ces hypothèses permettent d'écrire le résidu quadratique moyen sur un sous-domaine de taille  $k$  sous la forme  $\delta_k^2 = k \sigma_{b(N)}^2 = k \frac{\sigma_b^2}{N}$ . D'après l'équation (9.1.2), le *nombre de fausses alarmes* associé à une image de variance  $\sigma_{u(N)}^2$ , pour un sous-domaine de  $k$  pixels



contenu dans un domaine de  $n$  pixels, dépend du rapport  $\frac{\delta_k^2}{2\sigma^2}$  i.e., avec ces dernières hypothèses, du rapport

$$\frac{\delta_k^2}{2\sigma_{u(N)}^2} = \frac{k\sigma_b^2}{2\sigma_N^2} = \frac{k\sigma_b^2}{2\sigma^2}. \quad (9.4.29)$$

Par conséquent, dans le cas où les labels sont équidistribués, le facteur de résolution n'a pas d'influence sur le *nombre de fausses alarmes*.

En pratique, les classes sont en général très inégalement représentées dans une image en observation de la Terre (par exemple, eau très minoritaire, forêt majoritaire, etc.). L'hypothèse selon laquelle les classes sont équidistribuées n'est donc pas vérifiée habituellement et le mélange des classes peut être très variable d'un pixel à l'autre. Par la suite, nous tentons de définir l'évolution de la variance en fonction du facteur de résolution d'une manière plus réaliste, à partir de l'analyse empiriques d'images réelles.

Pour cette analyse, nous considérons, en tant qu'images de résolution maximale ( $u$ ), des images de fraction de couvert calculées sur des images **Spot/HRV** de la plaine du Danube. Rappelons que la fraction de couvert est un paramètre linéaire en fonction de la résolution. L'analyse est réalisée à partir de six images acquises à des dates différentes : les 15 et 31 Octobre 2000, le 22 Février, 18 Avril, 06 Juin et 05 Juillet 2001. Chaque image a été dégradée spatialement par moyennage par blocs de taille  $\sqrt{N} \times \sqrt{N}$ , pour  $\sqrt{N} \in \llbracket 1..80 \rrbracket$ . En effet, la variance d'une image d'une même scène peut être très variable d'une date à l'autre dans la mesure où le signal envoyé par un type d'occupation du sol peut être très différent selon les dates et particulièrement les saisons et, de plus, certaines classes peuvent apparaître ou disparaître au cours du temps faisant ainsi varier le mélange des classes au sein d'un pixel **BR**. La figure 9.2 présente l'évolution des écart-types de chaque image en fonction du facteur de résolution  $N$  considéré.

Remarquons que l'évolution des écart-types des différentes images est relativement similaire, avec un niveau plus ou moins élevé selon la date d'acquisition. Plus précisément, sur cet échantillon, la valeur de l'écart-type diminue à l'approche de l'hiver, jusqu'à environ 0.1 en Février, puis augmente à l'arrivée de l'été pour atteindre une valeur d'environ 0.27 en Juillet. Pour chaque image considéré, l'écart-type diminue de moins en moins lorsque le facteur de résolution  $N$  augmente. En effet, il diminue d'environ 0.05 entre le facteur  $N = 1$  et  $N = 20 \times 20$ , puis à peu près d'autant entre  $N = 20 \times 20$  et  $N = 80 \times 80$ .

Modéliser cette fonction nous permettrait de définir un modèle réaliste de la variance d'une image en fonction du facteur de résolution  $N$ , et de l'introduire dans l'expression du  $NFA$  (9.1.2). Le *nombre de fausses alarmes* adopterait alors vraisemblablement pour paramètre le facteur de résolution  $N$ , ce qui nous permettrait d'étudier son influence sur la détection.

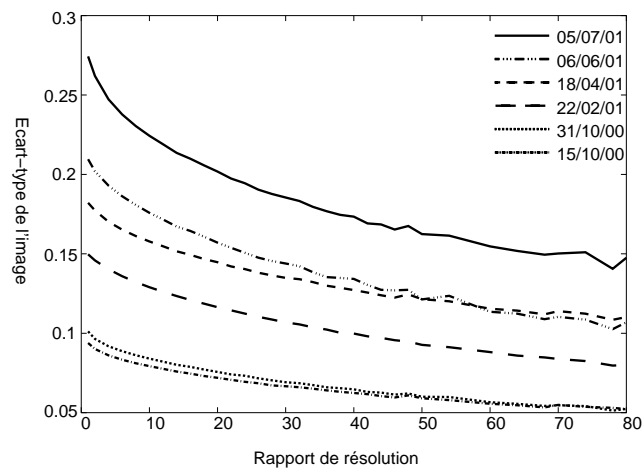


FIG. 9.2.: Evolution de l'écart-type en fonction du facteur de résolution. Ecart-type calculé sur des images de fraction de couvert de la série ADAM (échelle SPOT), et sur leurs dégradations spatiales pour des facteurs allant de 1 à 80. Les courbes sont relativement superposables. Image acquise le 15/10/2000 en marron (jaune), le 31/10/2000 en rouge, le 22/02/2001 en turquoise, le 18/04/2001 en vert, le 06/06/2001 en bleu et le 05/07/2001 en mauve.



## 10. Performances empiriques

L'objectif de ce chapitre est d'évaluer empiriquement les performances et limites de la détection en fonction des différents types de changements susceptibles d'apparaître dans une image BR. Ce chapitre présente plusieurs expériences réalisées à partir de données simulées. La partie IV sera ensuite consacrée à la présentation des résultats obtenus dans des cas réels d'applications.

### 10.1. Evaluation de la performance des résultats

La méthode décrite dans le chapitre 7 a été développée dans l'objectif de détecter un ensemble de pixels **BR cohérent** avec une labellisation **HR** donnée. Les pixels détectés par cette méthode doivent donc correspondre aux pixels de l'image qui sont sans changement par rapport à la date de référence à laquelle la classification a été réalisée. Le complémentaire de cet ensemble représente alors l'ensemble des pixels correspondants aux changements.

L'évaluation quantitative des performances est délicate dans le cas réel à partir du moment où elle dépend de la "vérité terrain", difficile à établir. De plus, les différents types d'erreurs de détection qui peuvent se produire n'ont pas nécessairement la même importance selon l'application. Pour s'affranchir du problème de la vérité terrain, l'analyse des performances est réalisée uniquement sur des données simulées (*cf.* section 10.2). L'évaluation quantitative des performances est alors réalisée en comparant les pixels détectés comme cohérents avec la labellisation (dits *structurés*) aux pixels correspondant effectivement à la labellisation d'après la vérité terrain. Plus précisément, il s'agit de compter le nombre de pixels détectés à tort et le nombre de détections manquées. D'un point de vue statistique, ces erreurs peuvent être vues comme des erreurs dites, respectivement, de type 1 et 2, en définissant l'hypothèse nulle du test par : 'le domaine  $D$  n'est pas structuré'. De manière plus générale, la méthode réalise une erreur de type 1 lorsqu'elle rejette l'hypothèse nulle à tort (faux positif), et une erreur de type 2 lorsqu'elle l'accepte à tort (faux négatif). Dans ce chapitre, notons qu'un pixel est dit détecté lorsqu'il est détecté par la méthode *a contrario*. Il ne s'agit donc pas d'un pixel représentant un changement mais, au contraire, d'un pixel cohérent avec la labellisation. L'analyse empirique des performances est fondée sur l'évaluation des quantités suivantes, très utilisées dans la littérature pour ce type d'étude.

**Définition 10.1.1 (Positifs)** *L'ensemble des positifs est l'ensemble des pixels détectés comme cohérents avec la labellisation de référence.*

<div style="text-align: right;">Vérité</div> <div style="text-align: left;">Décision</div>	sans changement (A+C)	avec changements (B+D)
sans changement (A+B)	vrais positifs (A) nombre de pixels structurés et détectés comme tels	faux positifs (B) nombre de pixels non structurés et détectés comme structurés à tort
avec changements (C+D)	faux négatifs (C)	vrais négatifs (D)

TAB. 10.1.: Récapitulatif des différents types d'erreurs.

Les pixels *positifs* correspondent donc aux pixels détectés *sans changements*. La définition suivante permet de distinguer, parmi les pixels positifs, ceux qui ont été détectés à tort de ceux qui ont été détectés à raison.

**Définition 10.1.2 (Vrais positifs)** *L'ensemble des vrais positifs est l'ensemble des pixels positifs qui sont, en réalité, cohérents avec la labellisation.*

Du point de vue de l'application, l'ensemble des vrais positifs correspond à l'ensemble des pixels détectés comme cohérents avec la labellisation (*i.e.* sans changements) et qui le sont en réalité. A l'inverse, les faux positifs correspondent aux pixels positifs *i.e.* détectés par le *NFA* comme des pixels *sans changement*, mais qui ne correspondent pas en réalité à la labellisation de référence. D'un point de vue thématique, les faux positifs sont généralement appelés *non-détections* puisqu'il s'agit des pixels de changement non détectés. De la même manière, nous définissons l'ensemble des négatifs.

**Définition 10.1.3 (Négatifs)** *L'ensemble des négatifs est l'ensemble des pixels non détectés, i.e. considérés comme non-cohérents avec la labellisation.*

Les pixels *négatifs* sont donc supposés correspondre aux pixels de changements. Dans l'ensemble des négatifs (pixels non-détectés), il faut distinguer ceux qui n'ont pas été détectés à raison (vrais négatifs) de ceux pour lesquels la détection a échoué (faux négatifs).

**Définition 10.1.4 (Vrais négatifs)** *L'ensemble des vrais négatifs est l'ensemble des pixels négatifs qui correspondent effectivement à des pixels de changements.*

A l'inverse, les faux négatifs correspondent aux pixels négatifs qui, en réalité, sont cohérents avec la labellisation et, par conséquent, auraient du être positifs. Le tableau 10.1 présente un récapitulatif de ces différents types d'erreurs.

La notion de faux négatifs (ou erreurs de type 2) et de faux positifs (erreurs de type 1) est très utilisée dans la littérature. Introduite par Neyman et Pearson [Neyman et Pearson, 1967] en 1930, elle est adoptée en statistiques mais aussi dans le cadre de nombreuses applications (analyses

médicales, reconnaissance de formes, recherche dans des bases de données, etc.). Le pourcentage de faux positifs ( $\frac{\text{nombre de faux positifs}}{\text{nombre total de positifs}} \times 100 = \frac{B}{A+B} \times 100$ ), par exemple, permet de quantifier le pourcentage de pixels détectés à tort par rapport au nombre total de détections. Cependant, il peut être intéressant d'analyser les différents types d'erreurs relativement à la réalité plutôt que relativement à l'ensemble des détections ou non-détections. En effet, par définition, les faux positifs font partie des pixels de changements, en réalité, et les faux négatifs font partie des pixels sans changements. Cette remarque nous invite à considérer le pourcentage de faux négatifs relativement à la vérité

$$\frac{\text{nombre de faux négatifs}}{\text{nombre réel de non-changements}} \times 100 = \frac{C}{A+C} \times 100 \quad (10.1.1)$$

et le pourcentage de faux positifs relativement à la vérité

$$\frac{\text{nombre de faux positifs}}{\text{nombre réel de changements}} \times 100 = \frac{B}{B+D} \times 100. \quad (10.1.2)$$

Dans ce chapitre, nous réalisons une série d'expériences à partir de données simulées comme décrit dans la section 10.2. Intuitivement, les deux principaux facteurs susceptibles de perturber la qualité de la détection sont, d'une part, la présence d'un trop grand nombre de pixels de changements dans l'image et, d'autre part, le taux d'occupation d'un changement dans le pixel **BR** concerné. Nous analysons l'évolution des différents types d'erreurs obtenus selon le nombre de pixels cohérents avec la labellisation (*sans changements*) qui sont présents dans l'image **BR** (section 10.3), puis selon la surface d'impact d'un changement dans un pixel **BR** dans la section 10.4.

Pour chaque expérience réalisée, nous présentons le pourcentage d'erreurs obtenues au total et en distinguant les faux positifs des faux négatifs, puis les pourcentages de faux négatifs relativement au nombre réel de pixels sans changements, et les pourcentages de faux positifs relativement au nombre réel de pixels de changements.

## 10.2. Simulation des données sans changements

Les images utilisées dans ce chapitre ont été simulées en utilisant des caractéristiques géométriques issues d'un extrait de taille  $256 \times 256$  d'une image **Spot/HRV** de la plaine du Danube<sup>1</sup>. Cet extrait a été segmenté en 100 régions par la méthode de [Mumford et Shah, 1989] à l'aide d'un algorithme de croissance de régions [Koepfler et al., 1994] (module **segct** du logiciel de traitement d'images MegaWave2<sup>2</sup>). Un label  $l \in \{0, 1, \dots, 9\}$  a ensuite été attribué aléatoirement à chaque région. L'image ainsi obtenue est appelée *labellisation* (c.f. figure 10.1), et sa résolution correspond à celle des images **HR** que nous utilisons.

A partir de cette labellisation, une image **HR** est simulée par tirage aléatoire, en chaque pixel, selon une loi normale paramétrée en fonction du label correspondant au pixel considéré. Une

<sup>1</sup>Image de la série ADAM (Assimilation de Données et Agro-Modélisation) mise à disposition par le CNES.

<sup>2</sup>Logiciel gratuit de traitement d'images, cf. <http://www.cmla.ens-cachan.fr/Cmla/Megawave/>

moyenne  $\mu_l$  est associée à chaque label  $l$ ,  $\mu_l \in \{0, 0.1, \dots, 0.9\}$ , et l'écart-type de chaque label est fixée en général à 0.05 (*i.e.* la demi-distance inter-classes). Une image obtenue de cette manière est présentée figure 10.1 pour les écarts-type  $\sigma = 0.05$ , et  $\sigma = 0.1$ .

L'image **BR** résulte du moyennage par blocs de taille  $N$  de l'image **HR**. Sous l'hypothèse selon laquelle les pixels sont des carrés adjacents, l'image **HR** est supposée prendre des valeurs constantes dans chaque carré  $1 \times 1$ , et l'image moyennée admet en chaque pixel la valeur moyenne de l'image **HR** obtenue sur le carré  $\sqrt{N} \times \sqrt{N}$  correspondant. La figure 10.1, à droite, présente une image **BR** simulée à partir d'une image **HR** obtenue par tirage gaussien d'écart-type 0.05 puis moyennage par blocs de taille  $16 \times 16$ .

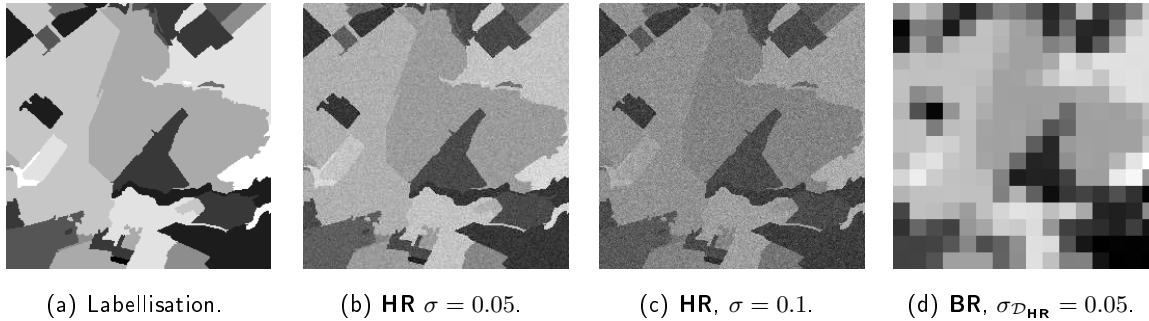


FIG. 10.1.: Images **HR** simulées par tirage gaussien à partir de la labellisation (a), avec un écart-type  $\sigma = 0.05$  (b) et  $\sigma = 0.1$  (c). L'image **BR** (d) est obtenue par moyennage de l'image **HR** (b) par blocs de taille  $16 \times 16$ .

Les images ainsi simulées sont, par construction, cohérentes avec la labellisation utilisée pour la simulation. Pour analyser les performances de la méthode de détection, nous proposons d'introduire des changements sur les images simulées.

### 10.3. Influence du nombre de pixels sans changement

Dans cette section, nous proposons de tester la robustesse de la méthode en fonction de la proportion de pixels touchés par des changements dans l'image **BR**. En général, les méthodes développées pour ce type d'applications ne permettent pas de détecter les zones de changements dès que le nombre total de pixels touchés par des changements concerne une part trop importante du nombre total de pixels de l'image. Par exemple, la méthode de [Le Hégarat-Masclé et Seltz, 2004] pour la détection de changements sur des images **BR** a tendance à détecter des changements sur tous les pixels de l'image dès que celle-ci contient plus de 20 à 30% de pixels de changements. Ce type de configuration se produit en réalité si les changements couvrant la totalité des pixels **BR** sont nombreux, mais aussi s'il manque une des données (par exemple, si le champ de vision du satellite n'est pas exactement identique d'une période à l'autre) ou si la valeur mesurée en certains pixels est aberrante (nuages, problèmes de capteur, etc.).

Dans cet objectif, nous commençons par présenter la manière dont les changements sont introduits sur les images simulées (section 10.3.1), puis nous décrivons les expériences réalisées et les résultats obtenus (section 10.3.2).

### 10.3.1. Simulation des changements

Le nombre de pixels utilisables dans une image peut nettement diminuer selon la position du satellite d'une acquisition à l'autre, ainsi que, par exemple, avec l'apparition de nuages couvrants, ou de changements de type de couvert affectant toute la surface d'un pixel basse résolution. De plus, selon les paramètres d'acquisition (date, heure, bande spectrale, etc.) et la scène étudiée, le contraste de l'image varie et son analyse est plus ou moins évidente. Dans le chapitre 9, nous avons mis en évidence l'importance du niveau de contraste de l'image pour la détection de changements, où le contraste est défini par le rapport signal sur bruit de l'image. Dans cette section, nous proposons d'analyser la performance de la méthode en fonction du nombre total de pixels sans changements présents dans l'image pour différents niveaux de contraste fixés, à partir d'images simulées comme décrit dans la section 10.2. Plus précisément, nous expérimentons la méthode sur des images **BR** résultantes du moyennage spatial d'images **HR** simulées comme décrit dans la section 10.2 avec les écart-types suivants :

- $\sigma = 0$ , *i.e.* image constante par morceaux (suivant la labellisation) à valeurs dans l'ensemble des valeurs moyennes  $\{0, 0.1, \dots, 0.9\}$ .
- $\sigma = 0.05$ , *i.e.* la demi-distance interclasses.
- $\sigma = 0.1$ , *i.e.* la distance interclasses.

Les écart-types nuls ou égaux à 0.1 correspondent typiquement à des cas limites pour l'application. En effet, les images présentent, en réalité, une variabilité à la fois interclasses et intraclasse. Cependant, le recouvrement des différentes classes n'est que partiel en général, c'est à dire que la variance interclasses est inférieure à la distance interclasses. Considérer ces différents écart-types pour la simulation revient finalement à considérer des niveaux de contraste forts ( $\sigma = 0$ ), moyens ( $\sigma = 0.05$ ) et faibles ( $\sigma = 0.1$ ).

Dans l'objectif d'analyser l'influence du nombre de pixels sans changements vis-à-vis des performances de la détection, nous proposons d'introduire des changements sur les images **BR** simulées avec ces différents niveaux de contraste et dans une proportion allant de 0 à 100% de l'image. Ces changements affectent l'intégralité des pixels basse résolution, ils sont simulés en dégradant par un bruit impulsionnel les images précédemment simulées avec différents niveaux de contraste. Plus précisément, il s'agit d'affecter des intensités aléatoires à  $p$  pourcents des pixels d'une image  $v$ , ce qui revient à transformer  $v$  en une image  $v'$  telle que, pour tout pixel  $y \in \mathcal{D}_{BR}$ ,

$$v'(y) = \begin{cases} b(y) & \text{si } n(y) = 1 \\ v(y) & \text{si } n(y) = 0, \end{cases} \quad (10.3.3)$$

où  $n$  est un champ de variables aléatoires de Bernoulli prenant la valeur 1 avec probabilité  $p$  est la valeur 0 avec probabilité  $1 - p$  et  $b$  est un champ de variables i.i.d. uniformément distribuées dans



l'intervalle  $[\min v, \max v]$ . La localisation des pixels *de changements* est donc tirée aléatoirement dans l'image, et la nouvelle valeur qui leur est attribuée est ainsi comprise entre les valeurs minimale et maximale de l'image. En particulier, il se peut qu'elle soit, par hasard, très proche de sa valeur initiale. Les pixels concernés risquent alors de ne pas être détectés. Par ailleurs, le pourcentage de pixels modifiés n'est pas toujours précisément égal à  $p$ . Cette variabilité se répercute directement sur la quantification de la qualité des résultats.

Dans la section 10.3.2, nous présentons les expériences réalisées et les résultats obtenus avec des images simulées de cette manière.

### 10.3.2. Résultats

Dans cette section, les performances de la méthode vis-à-vis du nombre total de pixels de changements présents dans l'image sont présentées pour des images simulées avec la résolution du capteur **Spot/HRV** pour la haute résolution et du capteur **Meris** pour la basse résolution comme indiqué dans la section 10.2. Le rapport de résolution que nous considérons dans cette partie est donc constant, égal à 16. Les changements sont simulés sur les images **BR** comme indiqué dans la section 10.3.1. Pour un pourcentage  $p$  de changements donnés, nous réalisons 500 tests à partir de 500 images représentant le même paysage mais contenant  $p\%$  de changements tirés aléatoirement. Pour nous affranchir du biais introduit par le tirage aléatoire des pixels et des valeurs qui leur sont attribuées lors de la simulation des changements, les résultats que nous présentons ici correspondent en général à la valeur médiane des résultats obtenus pour ces 500 expériences, ou la valeur moyenne lorsqu'elle est plus informative.

A partir d'une image **BR** simulée comme décrit dans la section 10.2 et de la labellisation (section 10.2), l'algorithme recherche le sous-domaine basse résolution qui minimise le nombre de fausses alarmes. Le seul paramètre laissé à l'utilisateur est le nombre d'itérations à réaliser. Dans la section 8.3, le problème de l'estimation du nombre d'itérations nécessaire est abordé, mais la complexité du problème ne permet pas d'en faire une estimation raisonnable. Cependant, rappelons que le nombre d'itérations doit être d'autant plus élevé que les pixels sans changements sont rares dans l'image. En effet, les bons pixels pour l'estimation des moyennes des classes sont alors rares, et la probabilité de les sélectionner par hasard devient très faible. En pratique, pour que les résultats soient comparables, nous proposons de réaliser chaque expérience avec un nombre fixe d'itérations (600 000), même s'il peut arriver que davantage d'itérations soient nécessaires pour mener à bien la détection, en particulier lorsque l'image contient en majorité des pixels de changements.

La figure 10.2 présente, en pourcentage du nombre de pixels de l'image, la médiane des erreurs obtenues pour ces 500 expériences en fonction du pourcentage de bruit impulsif (changements) introduit, et pour des images de différents niveaux de contraste, simulées avec un écart-type  $\sigma \in \{0, 0.05, 0.1\}$ . Chaque graphe représente, en fonction du pourcentage de bruit impulsif dans l'image, le pourcentage des erreurs obtenues au total (en rouge), le pourcentage des pixels de l'image détectés comme cohérents avec la classification à tort (faux positifs, en

bleu) et de ceux détectés comme changements à tort (faux négatifs, en vert). Remarquons, tout d'abord, que dans le cas plutôt réaliste des images de niveaux de contraste moyen ( $\sigma = 0.05$ ), contenant environ 20% de bruit impulsionnel, les résultats obtenus sur 500 tests (cf. figure 10.2 (b)) contiennent alors moins de 1.5% d'erreurs de détection, en médiane, *i.e.* sur 50% des tests réalisés, moins de 1.5% des pixels de l'image sont détectés à tort (faux positifs) ou non-détectés à tort (faux négatifs). En ce qui concerne le cas limite où le contraste est très faible ( $\sigma = 0.1$ ), nous observons environ 2.5% d'erreurs de détection en présence de 20% de bruit impulsionnel. Le pourcentage des erreurs observées est donc d'autant plus fort que le contraste est faible, il reste cependant très faible globalement. De plus, nous pouvons considérer, d'après la figure 10.2, que la détection est fiable même en présence de 70% de bruit impulsionnel où l'on compte moins de 10% d'erreurs même pour le niveau de contraste le plus faible ( $\sigma = 0.1$ ). Cette performance de 70% est à relier au nombre d'itérations effectuées, elle est d'autant plus élevée que le nombre d'itérations est élevé.

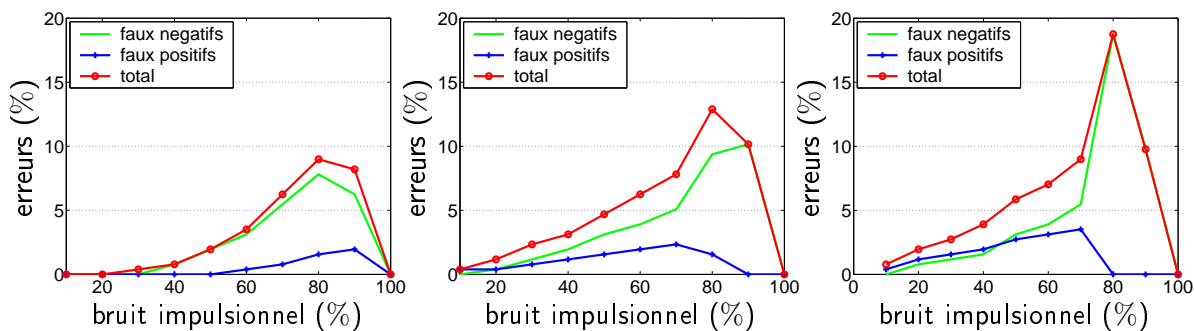


FIG. 10.2.: Pourcentage des pixels de l'image détectés à tort comme cohérents avec la labellisation (faux positifs, en bleu), de ceux détectés à tort comme changements (faux négatifs, en vert), et le total des erreurs obtenues en rouge. Valeur médiane des résultats obtenus sur 500 tests pour chaque pourcentage de bruit impulsionnel donné, avec des images simulées avec un écart-type  $\sigma = 0$  (a),  $\sigma = 0.05$  (b) et  $\sigma = 0.1$  (c). Même avec un faible niveau de contraste, la méthode permet de détecter les pixels de changement avec moins de 10% d'erreurs en présence de 70% de bruit impulsionnel.

Ces résultats montrent la robustesse de la méthode en présence de nombreux pixels de changements dans l'image. Cette propriété est due à la stratégie de tirage (RANSAC) adoptée pour l'algorithme de détection qui permet d'être très peu sensible à la présence de points aberrants lors de la phase d'estimation. Cet avantage est important pour notre étude dans la mesure où les images **BR** utilisées sont souvent incomplètes. En effet, dans le cas des images *grand champ*, il est fréquent que des nuages couvrent une partie de la zone de couverture, perturbant ainsi l'acquisition de la lumière réfléchie par le sol.

Globalement, l'évolution des erreurs de détection en fonction du pourcentage de bruit impulsionnel introduit est similaire pour les différents niveaux de contraste envisagés ( $\sigma \in \{0, 0.05, 0.1\}$ ), croissante jusqu'à près de 80% de pixels de changements dans l'image, puis fortement décroissante.

En effet, au-delà de 80% de changements, les chances de sélectionner par hasard un sous-système qui permette de bien estimer les caractéristiques des classes sont moins fréquentes pour un nombre d'itérations fixé (nombre de tirages). La méthode a alors tendance à sous-détecter les pixels sans changements, voire à ne détecter aucun pixel sans changement. Dans ce cas, le nombre de pixels positifs est nul, ainsi que le nombre de faux positifs. Cette décroissance nette ne révèle donc pas une amélioration des performances de la méthode mais simplement l'absence de toute détection. Si 80% des pixels de l'image représentent des changements et qu'aucune erreur n'est comptée, la décision prise pour 20% des pixels est alors erronée. Au-delà, la méthode ne détecte toujours pas de pixel valide, cependant le nombre de pixels valides est alors de plus en plus faible et, par conséquent, la décision prise est plus proche de la vérité.

En ce qui concerne la répartition des erreurs, entre détections et non-détections à tort (faux positifs et négatifs), elle évolue avec le niveau de contraste des images **BR** et la quantité de bruit impulsionnel introduit. En particulier, les faux négatifs sont largement prédominants pour les images à fort contraste ( $\sigma = 0$ ) dès que les changements concernent 40% de l'image **BR**. Lorsque le contraste diminue ( $\sigma = 0.05$  et  $\sigma = 0.1$ ), cette prédominance tend à disparaître. Par ailleurs, le nombre de faux positifs est relativement stable à la fois dans son évolution en fonction du pourcentage de bruit impulsionnel et en valeur. Cette stabilité est attendue dans la mesure où le principe du *NFA* est précisément de contrôler (minimiser) l'espérance du nombre de faux positifs (fausses alarmes pour le modèle *NFA*), *i.e.* l'espérance du nombre de non-changements détectés à tort.

Pour une analyse complémentaire des erreurs obtenues, nous représentons, figure 10.3 (a), le pourcentage des pixels détectés comme changements à tort (faux négatifs) relativement au nombre de pixels qui ne correspondent réellement pas à des changements ( $C/(A + C)$ , cf. tableau 10.1). Nous ne présentons pas ici la médiane des erreurs qui est peu informative, dans ce cas, en raison de la très faible présence de faux positifs et de faux négatifs. Les courbes représentent le pourcentage moyen d'erreurs sur les 500 tests réalisés, à partir d'images simulées avec un écart-type  $\sigma = 0$  en vert,  $\sigma = 0.05$  en rouge et  $\sigma = 0.1$  en bleu. Avec cette nouvelle normalisation, remarquons que l'évolution du pourcentage de faux négatifs est très peu sensible au contraste de l'image, en particulier lorsqu'il y a moins de 70% de bruit impulsionnel dans l'image. Au-delà, par exemple en présence de 70% de bruit impulsionnel, il reste 30% de pixels sans changements dans l'image **BR**. Parmi ces 30% de pixels sans changements, 20% sont détectés en moyenne comme correspondant à des changements d'après la figure 10.3 (a). De même, le pourcentage moyen des pixels détectés comme cohérents avec la labellisation à tort relativement au nombre de pixels qui sont réellement des changements est représenté, figure 10.3 (b), en fonction du pourcentage de pixels de changements introduits dans l'image **BR**. Ce type d'erreurs est globalement très faible, avec environ 3% d'erreurs pour une image de contraste moyen ( $\sigma = 0.05$ ), en présence de 10 à 70% de changements dans l'image. En effet, pour un niveau de contraste fixé, ce pourcentage est quasiment constant lorsque le bruit impulsionnel augmente, jusqu'à 70% environ. Au-delà, comme nous l'avons remarqué précédemment, le pourcentage de faux positifs chute car la méthode a alors tendance à sous-détecter ou à ne rien détecter. L'évolution du pourcentage de faux positifs

semble très stable à l'introduction de bruit impulsionnel, et reste très faible en valeur par rapport au pourcentage de faux négatifs. Cette propriété vient directement du choix de contrôler le nombre de fausses alarmes (*i.e.* faux positifs). Le nombre de faux positifs semble très robuste à la présence de bruit impulsionnel et, finalement, seul le niveau de contraste semble avoir un effet sur ce type d'erreurs. En effet, pour 40% de bruit impulsionnel, le pourcentage de faux positifs passe de moins de 0.5% en moyenne pour des images à fort contraste ( $\sigma = 0$ ) à environ 3% pour des images simulées avec un écart-type  $\sigma = 0.05$  et environ 5.5% pour des images à contraste très faible ( $\sigma = 0.1$ ). En revanche, le pourcentage de faux négatifs semble plus sensible à la présence de bruit impulsionnel et l'effet du niveau de contraste devient probablement négligeable devant la présence de nombreux pixels de changements.

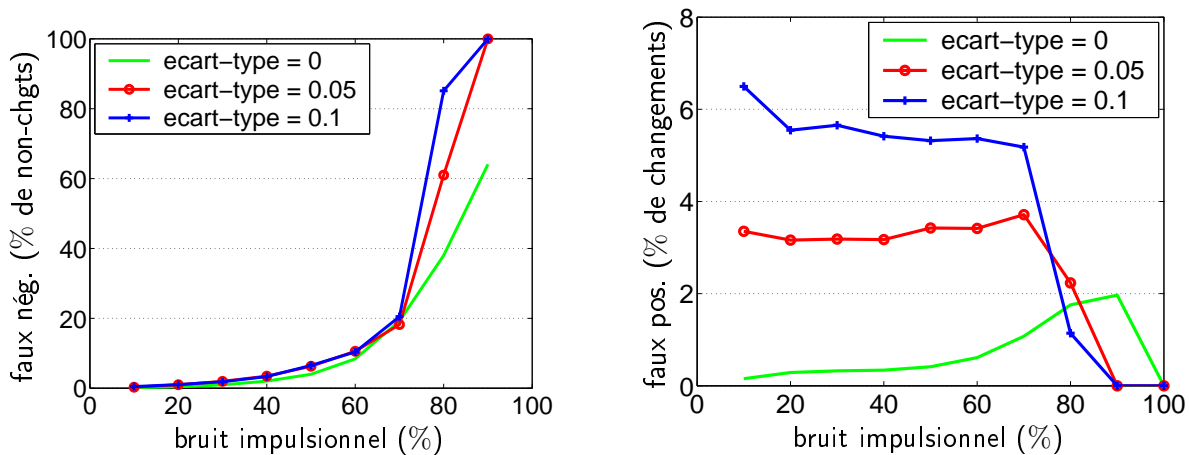


FIG. 10.3.: Erreurs de détection obtenues en moyenne en fonction du pourcentage de changements introduit. Le pourcentage moyen de pixels détectés comme changements à tort par rapport au nombre pixels sans changements est représenté figure (a) et, figure (b), le pourcentage moyen de pixels détectés comme cohérents avec la classification à tort par rapport au nombre pixels de changements. Le taux de faux négatif est très robuste au changement de contraste, même en présence de 75% de pixels de changement et le taux de faux positifs est très robuste au pourcentage de pixels de changement dans l'image (tant qu'il est inférieur à 75%).

La figure 10.4 présente les nombres des différents types d'erreurs obtenues en fonction du nombre de pixels de changements. Sur chaque graphe, le nombre effectif de changements est compté, et les résultats obtenus pour tous les tests réalisés sont représentés. Le graphe de gauche représente le nombre de vrais positifs (en vert), de faux négatifs (en bleu) et de pixels réellement sans changement (en rouge) en fonction du nombre de pixels de changements réellement présents dans l'image. Le nombre de pixels réellement sans changement est tracé simplement pour faciliter la lecture du graphe. Ainsi, plus le nuage de points des vrais positifs (verts) est proche de celui des pixels sans changements (rouge), meilleure est la qualité de la détection. De plus, rappelons que la somme des vrais positifs et des faux négatifs est égale au nombre de pixels réellement sans changements (*cf.* tableau 10.1). De même, le graphe de droite représente le nombre de vrais négatifs

(en vert), de faux positifs (en bleu) et de pixels correspondants réellement à des changements (en rouge) en fonction du nombre de pixels de changements réellement présents dans l'image. La détection des changements est alors d'autant plus performante que les points représentant les vrais négatifs sont proches de la première bissectrice. De plus, le nombre réel de changements est égal à la somme des vrais négatifs et des faux positifs. Ces deux graphes, complémentaires, sont présentés pour chaque niveau de contraste considéré ( $\sigma = 0$  fig. (a),  $\sigma = 0.05$  fig. (b) et  $\sigma = 0.1$  fig. (c)). Cette série de graphes permet d'observer clairement l'absence de toute détection dès que les pixels de changements deviennent trop présents dans l'image. Plus précisément, quelque soit le niveau de contraste des images testées, la méthode ne détecte plus rien dès que les changements concernent plus de 200 pixels environ, soit approximativement 80% de l'image **BR** comme nous l'avons déjà observé sur les figures 10.2 et 10.3. Par ailleurs, le fait que la méthode repose essentiellement sur le contrôle du nombre moyen de fausses alarmes (faux positifs) apparaît avec la comparaison entre les graphes de gauche et de droite (en bleu). En effet, le nombre de faux positifs (à droite) croît très peu lorsque le nombre de pixels de changement augmente alors que le nombre de faux négatifs a une croissance plus marquée en fonction du nombre de pixels de changements.

Pour finir, la figure 10.8 présente quelques exemples d'images **BR** simulées avec différents niveaux de contraste ( $\sigma \in \{0, 0.05, 0.1\}$ , respectivement figure 10.5, 10.6 et 10.7), et 20% de changements (a) ou 40% de changements (c) pour chaque niveau de contraste considéré. Pour plus de visibilité, l'image est présentée en composition colorée de sorte à faire apparaître les pixels sans changements en niveaux de gris et les pixels de changements en couleur. Les résultats de la détection obtenus pour les images (a) et (c) sont présentés, respectivement, figures (b) et (d) où les pixels détectés comme cohérents avec la labellisation à raison (vrais positifs) apparaissent avec leur valeur en niveaux de gris, ceux qui le sont à tort apparaissent en jaune, les pixels de changements détectés à raison (vrais négatifs) sont représentés en vert et à tort (faux négatifs) en rouge.

D'après les résultats obtenus sur des données simulées, la méthode semble particulièrement robuste face à la forte présence de points aberrants ou de pixels correspondants à des changements par rapport à la référence. En effet, les méthodes développées jusqu'à présent pour la détection de changements sont généralement limitées par la présence d'au moins 70% de pixels corrects. Cette limite est souvent étroitement liée à la méthode d'estimation utilisée, essentielle pour mener à bien la détection. L'utilisation d'une stratégie de type RANSAC couplée au critère du *nombre de fausses alarmes* constitue la pierre angulaire de notre méthode pour faire face à ce type de limites. En effet, cette stratégie est basée sur la résolution de sous-systèmes bien déterminés, sélectionnés aléatoirement, plutôt que sur l'utilisation de la totalité de l'échantillon pour estimer les paramètres. Même en présence de nombreux changements ou de points aberrants, un sous-système sélectionné peut, par hasard, ne contenir que des pixels sans changement. La qualité de la détection s'en trouve alors très peu altérée. Il suffit donc d'itérer le procédé suffisamment pour avoir au moins une chance de sélectionner un sous-système pertinent. Avec 600 000 itérations, les résultats restent très satisfaisants tant que 30% des pixels, au moins, sont corrects. La méthode permet alors une détection comportant moins de 8% d'erreurs sur des images de contraste moyen.

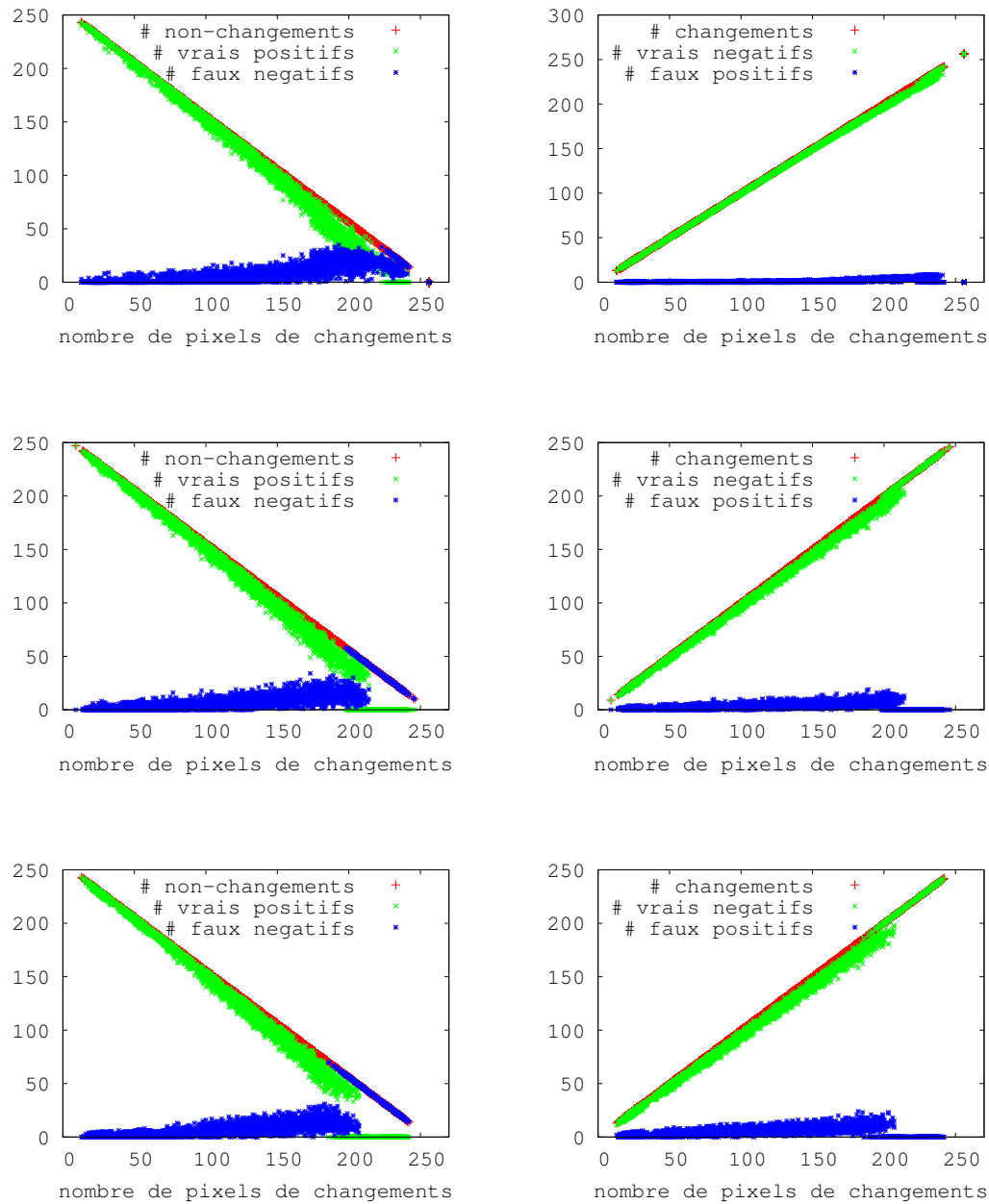


FIG. 10.4.: Evolution des performances en fonction du nombre de pixels de changements présents dans l'image. A gauche, les points rouge sur la diagonale représentent les pixels de non-changements présents en réalité dans l'image testée, le nombre de pixels détectés comme changements à tort (faux négatifs) est représenté en bleu, et le nombre de pixels détectés comme valides à raison (vrais positifs) en vert. A droite, les points rouge sur la diagonale représentent les pixels de changements, le nombre de pixels valides à tort est représenté en croix bleues et, en vert, le nombre de pixels détectés comme changements.

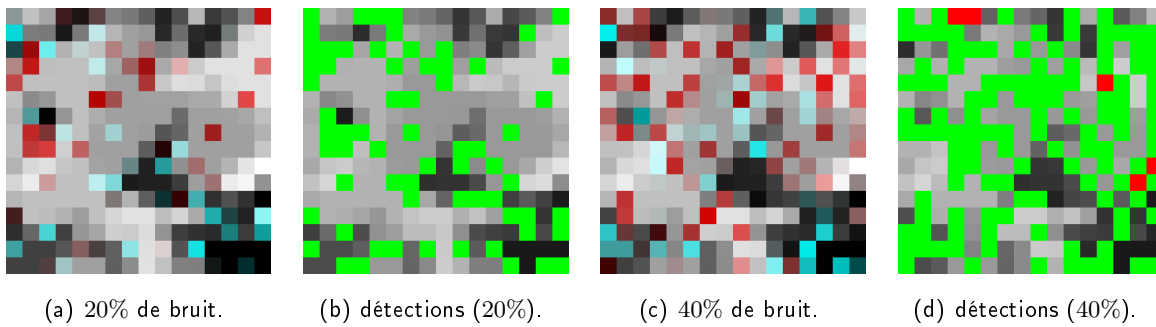


FIG. 10.5.: Images simulées avec un écart-type nul.

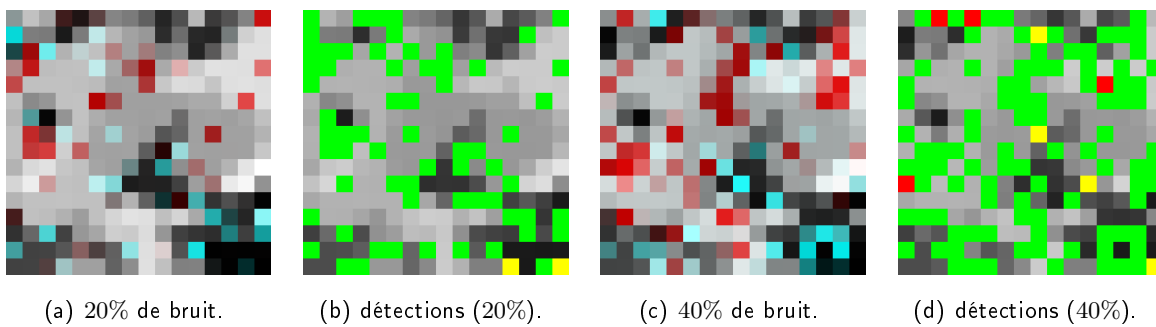


FIG. 10.6.: Images simulées avec un écart-type  $\sigma = 0.05$ .

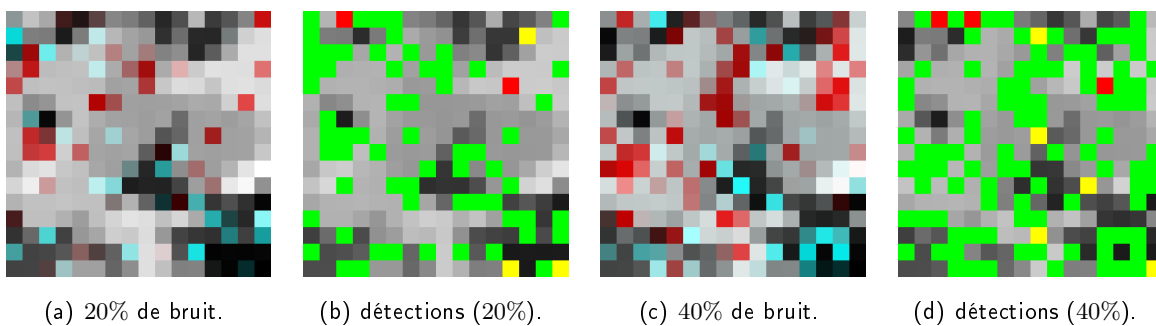


FIG. 10.7.: Images simulées avec un écart-type  $\sigma = 0.1$ .

FIG. 10.8.: Les images (a) et (c) font apparaître en rouge, respectivement, les 20% et 40% de bruit impulsif ajoutés à une image simulée avec la variabilité précisée dans la légende. Les pixels détectés par le NFA sont présentés figure (b) et (c) : les vrais négatifs en vert, les vrais positifs en niveaux de gris, les faux négatifs en rouge et les faux positifs en jaune. Globalement, les faux négatifs et les faux positifs apparaissent très peu.

Dans le contexte plus classique où environ 20% des pixels de l'image correspondent à des changements, la détection à partir d'une image de contraste moyen ( $\sigma = 0.05$ ) se fait avec moins de 1.5% d'erreurs au total, moins de 2.5% d'erreurs lorsque le contraste de l'image est faible ( $\sigma = 0.1$ ) et sans erreur s'il est fort. Même si ce taux d'erreurs est globalement très faible, nous observons qu'il augmente lorsque le contraste diminue, comme nous l'avons remarqué théoriquement avec la proposition 9.2.2 (section 9.2).

Dans la section 10.4, nous abordons le problème de la détection des changements sous-pixeliques, *i.e.* des changements de taille inférieure à celle du pixel basse résolution.

## 10.4. Influence de la taille du changement sous-pixelique

Après avoir analysé les performances de la méthode en fonction du pourcentage total de changements apparaissant dans l'image, nous proposons, dans cette section, d'étudier les performances obtenues lorsque les changements introduits n'affectent que partiellement un pixel basse résolution. Il s'agit d'un cas de figure très fréquent qui constitue une des principales difficultés de la détection de changements à partir d'images **BR**. Ces changements, dits sous-pixeliques, apparaissent dès que la taille des changements à détecter n'est pas suffisamment grande relativement à la taille d'un pixel **BR**, *i.e.* typiquement lorsqu'un changement se produit sur une parcelle agricole qui n'est que partiellement couverte par un pixel basse résolution.

Dans cette section, notre objectif est d'analyser les performances du détecteur en fonction du taux d'occupation d'un changement dans un pixel basse résolution. En particulier, la détection est-elle d'autant plus performante que le changement occupe une grande proportion du pixel basse résolution ? Existe-t'il un seuil en-dessous duquel les changements ne peuvent plus être détectés ? Pour tenter de répondre à ces questions, nous simulons un certain nombre d'images tests contenant des changements présents en des proportions données. La section 10.4.1 décrit le processus de simulation des images, puis la section 10.4.2 présente les résultats obtenus avec la méthode de détection *a contrario*.

### 10.4.1. Simulation des changements

En utilisant des images provenant de capteurs tels que **Spot/VGT** pour lesquels chaque pixel représente une surface au sol de  $1km^2$ , de nombreux phénomènes peuvent se produire sur une surface très inférieure, mais aussi sur une surface qui, bien que supérieure, se trouve projetée sur plusieurs pixels basse résolution en proportion éventuellement faible. Pour séparer les problèmes, nous proposons de nous placer dans un contexte où le nombre de pixels touchés par des changements est fixe, et représente 20% des pixels de l'image **BR**. Ce pourcentage a été choisi arbitrairement comme un pourcentage moyen pour le type d'application visée. De plus, la section 10.3 a permis d'observer un très bon fonctionnement de la méthode pour les images contenant 20% de pixels de changement, donc ce choix semble bien indiqué pour éviter de confondre les différentes limites



de la méthode tout en envisageant un cadre réaliste. Pour les mêmes raisons, nous considérons un niveau de contraste fixe, simulé à haute résolution avec un écart-type  $\sigma = 0.05$  correspondant à la demi-distance interclasses.

Par définition, les changements sous-pixelliques doivent être introduits sur les images haute résolution avant leur dégradation spatiale. Pour faciliter l'analyse, tous les changements introduits dans une image donnée concernent la même proportion d'un pixel **BR**. Par conséquent, nous simulons des images pour lesquelles tous les changements ne représentent que  $p\%$  de la surface d'un pixel **BR**, où  $p$  varie de 1 à 100. Dans cet objectif, nous sélectionnons aléatoirement 20% des pixels de l'image **BR** pour les modifier, et nous considérons leur projection sur le domaine haute résolution. Un changement impactant  $p\%$  d'un pixel **BR** est simulé à partir du tirage aléatoire uniforme d'un label  $l' \in \{0, \dots, 10\}$  correspondant au type d'occupation du sol du changement. Ce tirage permet de représenter un simple changement de label (parmi les labels existants) ou l'apparition d'un nouveau label (le label 10 n'est pas présent dans la labellisation initiale). Les valeurs des pixels **HR** de changements sont alors remplacées par les valeurs tirées aléatoirement selon une loi normale de moyenne  $0.1 * l'$  et d'écart-type 0.05. Ainsi simulé, un changement qui se produit sur la surface représentée par un pixel **BR** ne correspond *a priori* qu'à un seul type de perturbation (par exemple : coupe de forêt, inondation, incendies, conversion de prairie en zone agricole, nuages, etc.). Cependant, lorsque le label d'origine du pixel **HR** considéré est égal à  $l'$ , la valeur du label de changement  $l'$  est tirée aléatoirement une nouvelle fois. La nouvelle image **HR** est ensuite moyennée par blocs de taille  $16 \times 16$  de sorte à créer l'image **BR** test.

Dans la section 10.4.2, nous présentons les expériences menées et les résultats obtenus à partir d'images simulées avec des changements dont la taille varie relativement à la taille du pixel **BR**, mais le nombre total de pixels de changements est fixé.

### 10.4.2. Résultats

Dans l'objectif d'analyser les performances de la méthode en fonction de la surface que représente un changement au sein d'un pixel **BR**, nous considérons un rapport de résolution constant entre la labellisation et l'image **BR**, égal à 16. Ce rapport correspond typiquement au cadre d'application aux données provenant des capteurs **Spot/HRV** et **MeRIS**. Dans cette section, nous présentons les résultats de la détection obtenus à partir de la labellisation (cf. figure 10.1) et d'images simulées comme indiqué dans la section 10.4.1, où 20% des pixels de l'image **BR** sont touchés par des changements en proportion  $p$ . Pour tout  $p \in \{1, \dots, 100\}$  fixé, nous appliquons la méthode de détection sur 500 images simulées, contenant 20% de pixels affectés par des changements en proportion  $p$ . Pour éviter de biaiser les résultats, chacune de ces expériences a été réalisée avec un nombre fixe d'itérations égal à 400 000, nombre largement suffisant lorsque les changements ne concernent que 20% des pixels de l'image.

La figure 10.18 présente les pourcentage des différents types d'erreurs de détection obtenus en fonction du pourcentage de la surface d'impact d'un changement sur le pixel **BR** cor-

respondant. Pour plus de lisibilité, chacun de ces graphes présente la médiane et les 25<sup>eme</sup> et 75<sup>eme</sup> percentiles des erreurs obtenues sur tous les tests réalisés pour un taux d'occupation  $p$  du changement dans le pixel **BR** donné. De plus, les résultats ont été rassemblés par pas de 5, *i.e.* la médiane et les percentiles sont calculés pour tous les résultats obtenus avec  $p \in ]k, k + 5]$  pour tout  $k \in \{0, \dots, 95\}$ . Le pourcentage des erreurs de détection obtenues globalement est présenté figure 10.18, à gauche. D'après ce graphe, remarquons que la méthode est capable de détecter les changements avec moins de 5% d'erreurs (en médiane) dès que leur surface d'impact représente plus de 13% environ du pixel **BR**, et avec moins de 2.5% d'erreurs lorsque les changements représentent plus de 25% du pixel **BR**. Contrairement à ce qu'on pourrait penser intuitivement, les performances de la détection ne s'améliorent que très peu au-delà de ce seuil.

Pour une analyse plus complète des erreurs de détection, la figure 10.18, au centre et à droite, présentent respectivement les pourcentages de faux positifs et de faux négatifs obtenus relativement au nombre total de pixels de l'image **BR**. Le nombre de pixels touchés par des changements étant fixé pour ces expériences (à 20% des pixels de l'image), une normalisation sur le nombre réel de changements ne changerait pas l'allure des courbes. En particulier, remarquons que le seuil de 25% observé globalement pour les erreurs de détection correspond à une valeur seuil à la fois pour les faux positifs et pour les faux négatifs. Lorsque les zones de changements représentent moins de 10% des pixels **BR**, la méthode a tendance à détecter presque tous les pixels de l'image comme cohérents avec la labellisation. En réalité, 20% des pixels de l'image représentent des changements. Ils sont donc détectés à tort par la méthode, d'où la présence de près de 100% de faux positifs (*cf.* figure 10.18 au centre) lorsque  $p < 10\%$ . De plus, si presque 100% des pixels sont détectés par la méthode, le nombre de négatifs est presque nul. La normalisation sur le nombre réel de pixels sans changements étant élevé, le pourcentage de faux négatif (*cf.* figure 10.18 à droite) apparaît comme nul tant que  $p \leq 25\%$ . Au-delà de ce taux d'occupation, le pourcentage de faux négatifs est, en médiane, alterne entre 0 et 0.5%. C'est à dire que, dès que les changements occupent plus de 25% du pixel **BR**, seuls 0.5% des pixels (valeur médiane) réellement sans changements sont détectés à tort par la méthode comme correspondant à des changements.

Naturellement, la question se pose alors de savoir si les performances obtenues en fonction du taux d'occupation des changements dans chaque pixel **BR** impacté sont indépendantes du rapport de résolution considéré. Pour tenter de répondre à cette question, nous proposons de réaliser la même expérience que précédemment, avec les rapports de résolution 4, 16 et 32. Pour que les résultats soient comparables, les images **HR** ont été simulées en conséquence, *i.e.* de manière à avoir un niveau de bruit comparable, une fois l'image dégradée. En effet, une diminution de la résolution d'une acquisition se traduit par un effet de régularisation sur l'image, *i.e.* une diminution de sa variance empirique. Plus précisément, en supposant l'indépendance spatiale des données, nous pouvons supposer qu'une image **HR** de variance  $\sigma^2$  moyennée d'un facteur  $N$  admet alors pour variance  $\sigma^2/N^2$ . L'image **HR** est donc simulée respectivement avec les écart-types 0.0125, 0.05 et 0.1 respectivement pour les rapports de résolution 4, 16 et 32. Comme précédemment, nous fixons le pourcentage de pixels de l'image touchés par des changements à 20% et les changements

sont ensuite introduits sur les images bruitées de sorte à représenter de 0 à 100% de chaque pixel de changements. La figure 10.18 présente les différentes erreurs obtenues avec, de gauche à droite, le pourcentage total d'erreurs, de faux positifs relativement au nombre réel de pixels de changements et de faux négatifs relativement au nombre réel de pixels sans changements. Les trois courbes correspondant au trois facteurs de résolution considérés évoluent de manière très similaire en fonction du taux d'occupation des changements dans chaque pixel. En particulier, remarquons que la valeur seuil de 25% d'occupation des pixels est fixe pour ces trois rapports de résolution, quelque soit le type d'erreurs considéré. Cette expérience ne correspond qu'à un exemple particulier. Cependant, elle semble clairement indiquer la dépendance des performances au taux d'occupation des changements à l'intérieur des pixels **BR** plutôt qu'au rapport de résolution en lui-même.

La capacité d'une méthode à détecter les changements sous-pixeliques est un point clef pour la détection de changements à partir d'images **BR**. Globalement, d'après les expériences réalisées, les performances de la méthode sont stationnaires vis-à-vis du taux d'occupation d'un changement dans un pixel **BR** dès que ce taux est supérieur à 25%. Elles présentent alors moins de 2.5% d'erreurs de détection au total (valeur médiane) quelque soit le rapport de résolution considéré. Lorsque la taille relative des changements dans le pixel **BR** est inférieure à 25%, les résultats s'améliorent avec la taille des changements. Cependant, même lorsque les changements n'occupent que 5% des pixels **BR**, les erreurs de détection ne concernent qu'environ 10% des pixels (valeur médiane). De plus, les performances ne semblent dépendre du rapport de résolution qu'à travers le pourcentage d'occupation des changements dans le pixel **BR**.

Dans la partie IV, nous présentons quelques cas d'application des méthodes de classification et de détection de changements sur un jeu de données réelles de la plaine du Danube, en Roumanie.

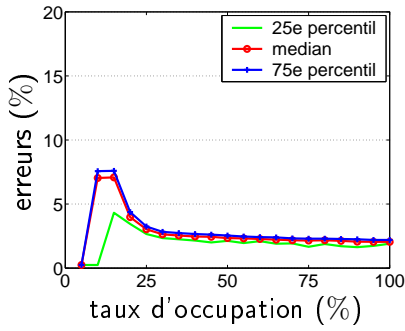


FIG. 10.9.: Erreurs (total)

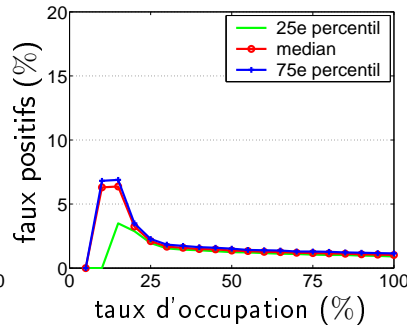


FIG. 10.10.: Non détections

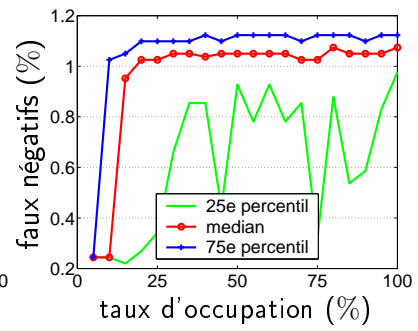


FIG. 10.11.: Non détections

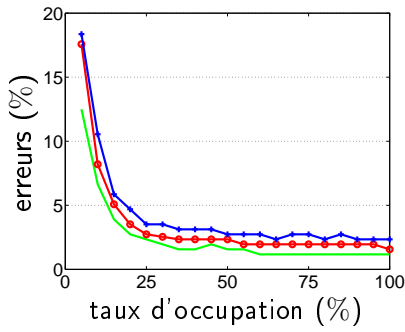


FIG. 10.12.: Erreurs (total)

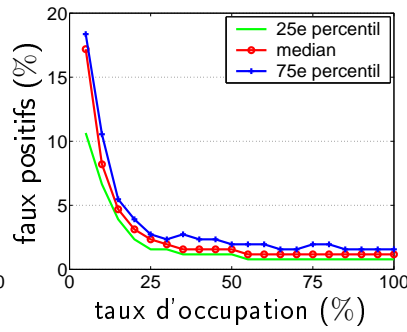


FIG. 10.13.: Non détections

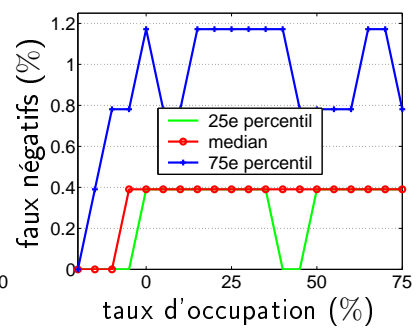


FIG. 10.14.: Non détections

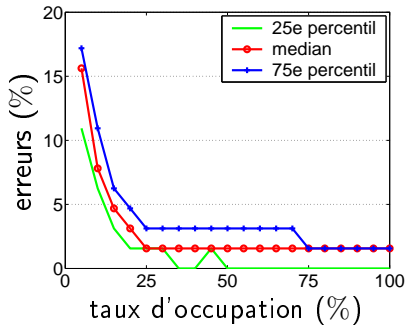


FIG. 10.15.: Erreurs (total)

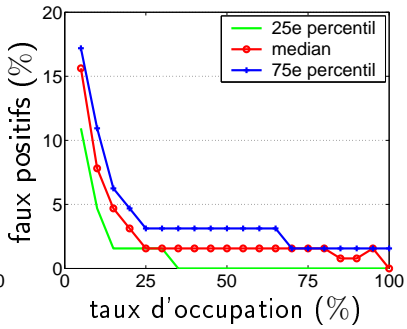


FIG. 10.16.: Non détections

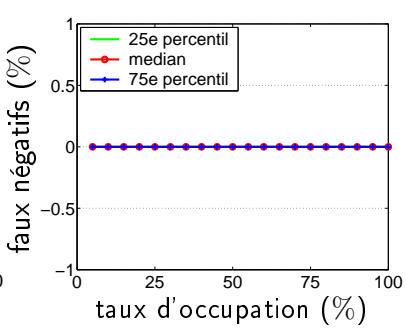


FIG. 10.17.: Non détections

FIG. 10.18.: Erreurs de détection (% du nombre de pixels dans l'image) obtenues en fonction du taux d'occupation des changements dans un pixel **BR**. De gauche à droite, médiane et percentiles du total des erreurs obtenues, des faux positifs et des faux négatifs. Performances obtenues pour des images simulées avec 20% de pixels de changements et, de haut en bas, des rapports de résolution de 4, 16 et 32.



Quatrième partie .

Application à une zone agricole de  
la plaine du Danube (base de  
données ADAM)



# 11. Introduction

## 11.1. La base de données ADAM

Dans cette partie, nous présentons quelques scénarii d'applications à la base de données ADAM. Celle-ci réunit, notamment, une série de 39 images **SPOT/HRV** acquises en mode multi-spectral (pixels représentant environ  $20 \times 20\text{m}^2$ ) entre octobre 2000 et juillet 2001 dans l'objectif d'évaluer l'apport des techniques d'assimilation de données spatiales pour la modélisation agronomique et l'information agricole en général. Le projet ADAM (Assimilation de Données par Agro-Modélisation) (*cf* <http://kalideos.cnes.fr/>) s'est intéressé à une zone d'agriculture intensive dans la plaine du Danube, en Roumanie. Ce site est caractérisé par un grand parcellaire, qui s'étend sur  $40 \times 60\text{ km}^2$  à environ et se trouve à 35 km à l'est de Bucarest. Dans ce cadre, les images **SPOT/HRV** ont été acquises environ une fois par quinzaine. Une telle fréquence d'acquisition (39 images en 10 mois) a été possible par l'utilisation conjointe des satellites **SPOT 1, 2** et **4**. Cependant, la présence de nuages ou de neige rend inutilisable certaines zones de l'image selon les dates et la qualité des images issues des différents satellites est variable.

De plus, un certain nombre de mesures ont été réalisées sur le terrain : caractéristiques du sol, du couvert, recensement des pratiques culturales. Ces mesures ont été réalisées sur 42 unités expérimentales couvrant chacune plus d'un pixel **SPOT** (environ 30m de diamètre) (*cf.* figure 11.1). Dans l'optique de valider les résultats obtenus, nous considérons principalement une zone d'étude restreinte à un carré  $256 \times 256$  autour des mesures de terrain présentées figure 11.1.

Après une série de pré-traitements réalisés par le CNES, les images utilisées sont superposables avec une précision inférieure au pixel (en projection GaussKruger) et correspondent à des mesures TOC (*Top of Canopy*), *i.e.* corrigées (en théorie) des effets de l'atmosphère. Les réflectances mesurées dans les trois canaux du capteur **SPOT/HRV** fournissent des informations redondantes pour l'étude de la végétation, dont les plus importantes proviennent des bandes 2 et 3 (*cf.* figure 11.2). De nombreux indices de végétation (*cf.* Section 1.3.2) sont définis à partir de ces deux bandes spectrales de manière à sélectionner l'information pertinente pour la végétation. Dans cette partie, la plupart des résultats présentés ont été obtenus à partir d'images de fraction de couvert (*cf.* figure 11.3). Ces images ont été générées à partir des trois canaux **SPOT/HRV** et des informations d'acquisition grâce à une chaîne de traitements<sup>1</sup> fondée sur le modèle biophysique SAIL + PROSPECT (*cf.* [Jacquemoud et al., 1995]) pour l'agriculture de précision.

---

<sup>1</sup>Chaîne de traitement mise à disposition par EADS/ASTRIUM.



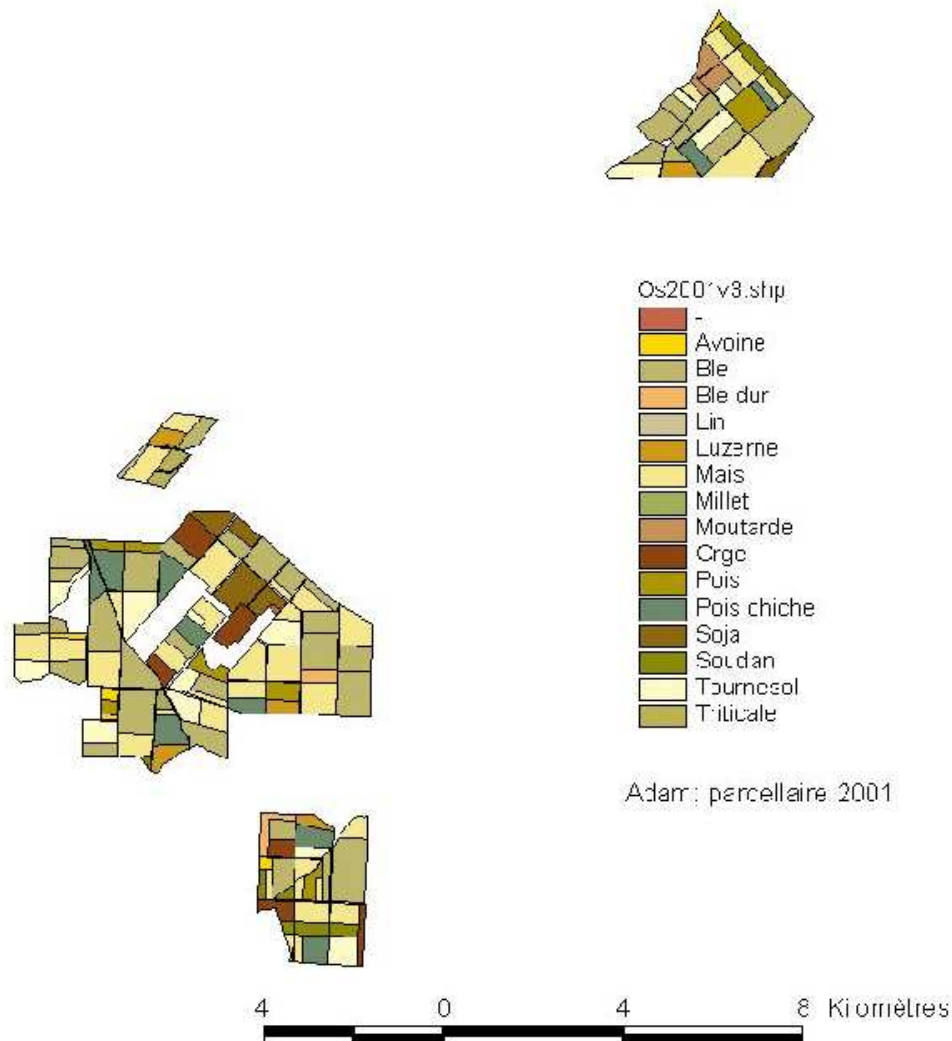


FIG. 11.1.: Bilan des mesures de terrain pour l'année agricole 2000 – 2001.

L'analyse de sensibilité présentée dans le chapitre 6 a montré, sur des données simulées, que les performances de la méthode de classification non-supervisée se dégradaient lorsque le rapport de résolution dépassait  $20 \times 20$ . Par conséquent, nous nous plaçons dans le cadre de l'utilisation conjointe d'images **SPOT/HRV** et **MeRIS** (*i.e.* dans un rapport de résolution  $15 \times 15$ ). Les seules images optiques **BR** de la base de données ADAM étant des images **SPOT/VGT** (*i.e.* dans un rapport de résolution  $50 \times 50$ ), les images **BR** ont été simulées à partir des données **SPOT/HRV** par moyennage par blocs de taille  $16 \times 16$ .

Dans la section 11.2, nous présentons quelques scénarii d'application considérés de manière à présenter les résultats obtenus dans différents contextes réalistes.

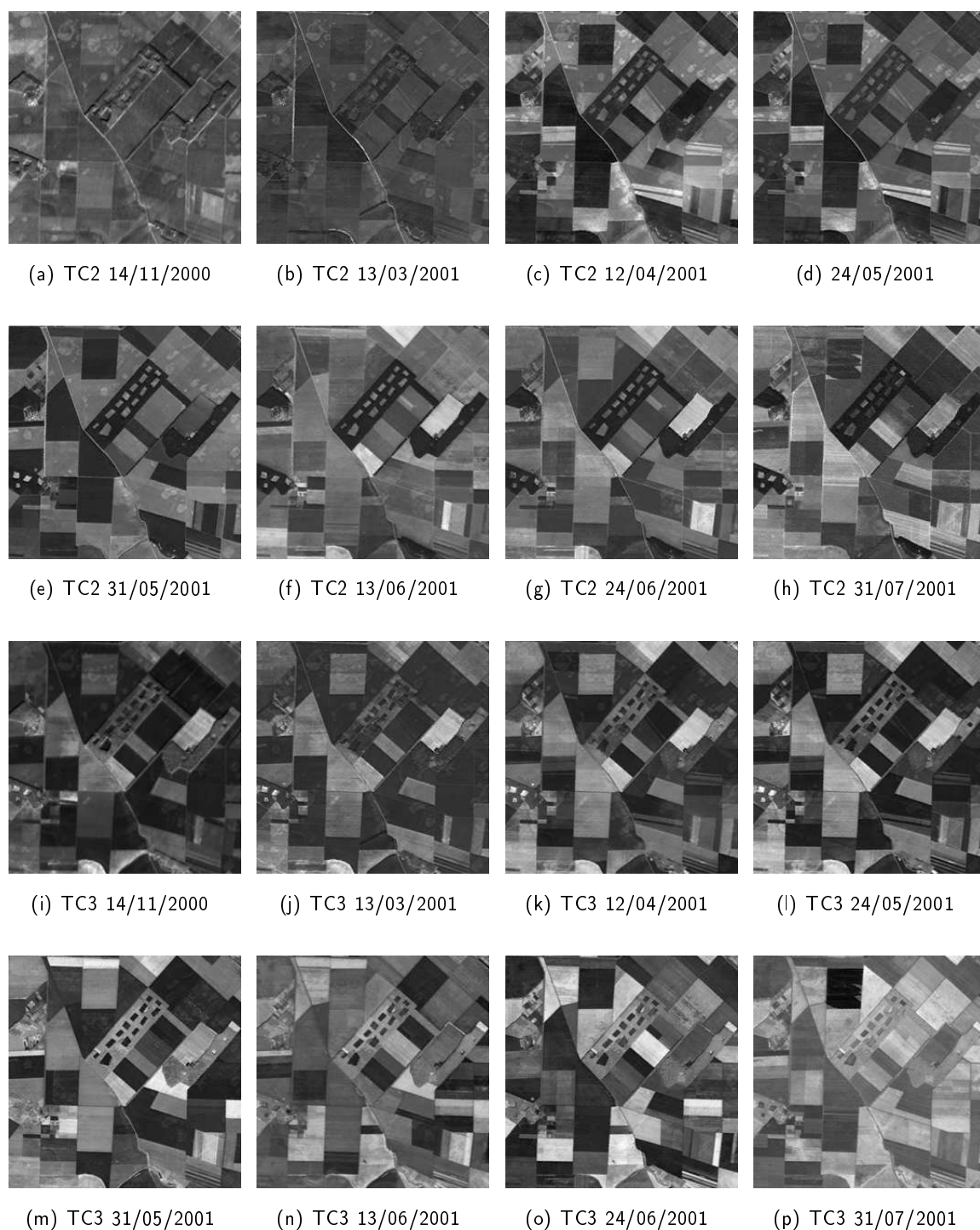


FIG. 11.2.: Images **HR** de réflectance acquises à 8 dates différentes, réparties entre le 14 novembre 2000 et le 31 juillet 2001. Les images acquises dans le canal 2 du capteur **SPOT/HRV** sont présentées en haut (TC2) et les images acquises dans le canal 3 sont présentées en bas (TC3).

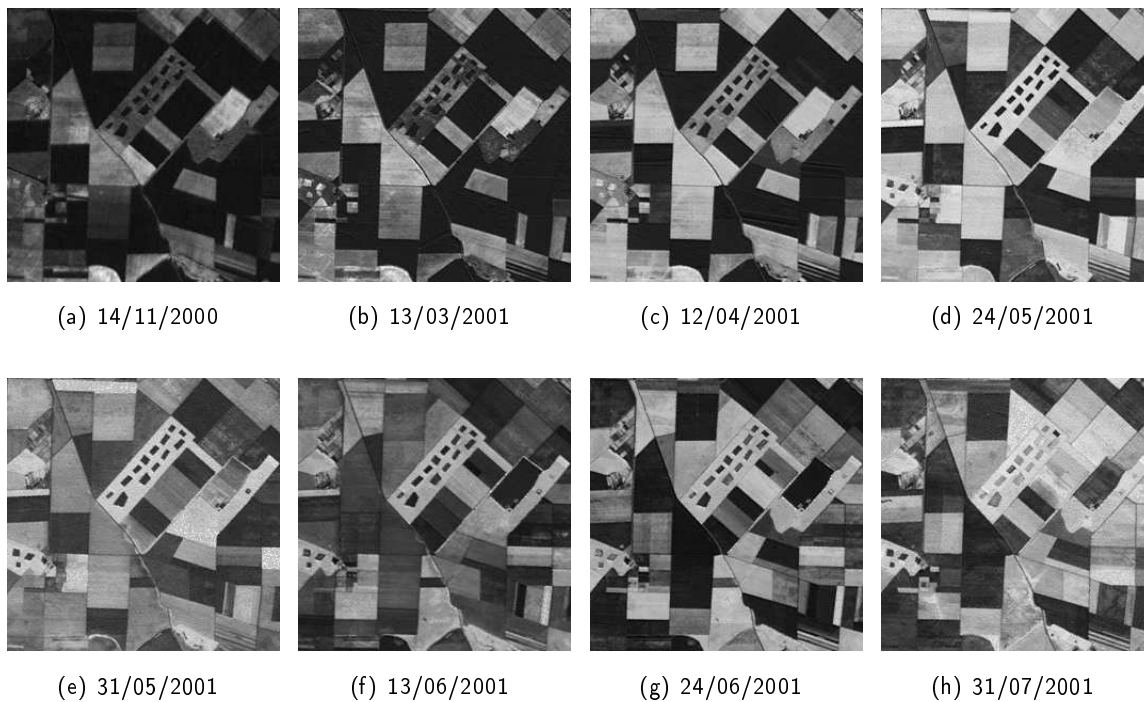


FIG. 11.3.: Images **HR** de fraction de couvert acquises à 8 dates différentes, réparties entre le 14 novembre 2000 et le 31 juillet 2001.

## 11.2. Quelques scenarii d'application

La problématique considérée au cours de ce travail de thèse a été définie, en particulier, pour l'analyse et le suivi de la végétation. Si les méthodes élaborées dans les parties II et III ont des domaines d'application distincts, leur utilisation conjointe conduit à un processus complet de mise à jour.

En effet, la méthode de classification sous-pixellique définie dans la partie II permet la création de cartes d'occupation des sols à **HR** à partir d'une segmentation **HR** et d'une séquence d'images **BR**. Cette méthode permet d'exploiter l'information radiométrique multitemporelle pour distinguer les différents types d'occupation du sol tout en bénéficiant de la localisation spatiale **HR** donnée par le parcellaire. L'ensemble de la surface terrestre pourrait alors être régulièrement cartographié à un moindre coût, permettant ainsi un suivi de l'occupation des sols pour des applications géographiques, agronomiques, ou politiques. Notons, cependant, que l'algorithme ne fournit pas directement une carte d'occupation des sols mais une carte de labels. Une interprétation thématique des différents labels est ensuite nécessaire pour obtenir une carte d'occupation des sols.

La méthode de détection de changements permet, quant à elle, de détecter des zones de changements d'occupation du sol même lorsqu'ils ne concernent qu'une partie du pixel **BR**. Plus précisément, cette méthode permet d'étudier la *cohérence* entre une séquence d'images **BR** et une classification **HR** décrivant l'état de la surface à une date de référence. Par conséquent,

si le critère défini permet de détecter la présence de changements avec une certaine robustesse (*cf.* chapitres 9 et 10), il ne permet pas de qualifier le changement ni de déterminer la surface impactée. En effet, une petite surface touchée par un fort changement d'intensité radiométrique peut être détectée de la même manière qu'un changement de moindre intensité mais ayant une surface d'impact supérieure.

Si les deux méthodes proposées sont indépendantes d'un point de vue méthodologique, elles peuvent être utilisées en collaboration pour la mise en œuvre de processus plus complets vis-à-vis de certaines applications. Par la suite, nous proposons trois scenarii d'application :

- lorsque des changements sont détectés, la méthode de classification peut être utilisée pour générer une nouvelle carte d'occupation du sol et, ainsi, déterminer le type de changement observé ;
- la détection du domaine **BR** le plus *cohérent* avec une carte de référence peut être exploitée dans l'objectif de valider une carte d'occupation des sols (ou classification) ;
- une carte d'occupation des sols peut être mise à jour par l'utilisation successive des méthodes de classification et de détection de changements : la séquence d'images **BR** est utilisée pour générer une classification, validée par la méthode de détection. Cette dernière permet ensuite, à l'acquisition de nouvelles images, de détecter si la carte est toujours valide ou si des changements sont apparus et la classification à régénérer.

Les chapitres 12 et 13 présentent les résultats obtenus dans le contexte de ces différents scenarii pour la zone d'étude du projet ADAM sur laquelle nous disposons de mesures de terrain.



## 12. Classification et validation

### 12.1. Classification sous-pixellique

Dans cette section, nous appliquons la méthode de classification non-supervisée (cf Partie II) à un extrait de la zone d'étude du projet ADAM pour lequel des relevés ont été effectués sur le terrain. Nous proposons d'évaluer les résultats de la méthode de classification non-supervisée d'un point de vue méthodologique et quantitatif en comparant les résultats obtenus à partir d'images **BR** à ceux obtenus à partir d'images **HR**, et d'un point de vue thématique et qualitatif en comparant les classifications que nous obtenons à la classification réalisée par F. Oro (INRA) à partir des relevés de terrain (figure 12.5) dans le cadre du projet ADAM.

L'algorithme de classification non-supervisé permet, étant donné un nombre de labels, de déterminer le label de chaque segment à partir d'une image ou d'une séquence d'images. La figure 12.1 présente les résultats de l'algorithme de labellisation non-supervisé obtenus à partir d'une segmentation **HR** comportant 100 régions (figure 12.1 (a)) et d'une séquence **HR** de huit images de fraction de couvert (figure 12.1 (b)) pour regrouper les 100 régions en 5 labels. Cette labellisation, générée à partir d'images **HR**, est considérée comme labellisation de référence pour évaluer les résultats obtenus à partir d'images **BR**. A partir de la séquence d'images **BR** de fraction de couvert 12.2 (a) et pour 5 classes, l'algorithme permet d'obtenir la labellisation 12.2 (b). Cette dernière est identique à la labellisation 12.1(c) pour 96.6% des pixels. L'image 12.2 (c) permet de visualiser, en blanc, les erreurs obtenues à partir de la séquence **BR** par rapport au résultat **HR**.

A titre de comparaison, la même expérience est réalisée sur la séquence d'images de  $NDVI^1$  correspondant aux mêmes dates (figure 12.3 (a)). L'objectif de cette comparaison n'est pas tellement d'évaluer le meilleur *indice* pour la classification des types d'occupation du sol mais plutôt de vérifier que l'utilisation de sources différentes conduit au même résultat. En effet, les images de  $NDVI$  ont été générées par combinaison des canaux 2 et 3 d'acquisitions **SPOT/HRVIR** alors que les images de fraction de couvert résultent d'une inversion complexe des mesures **SPOT/HRVIR** à travers un modèle biophysique. La labellisation (cf. 12.3 (b)) obtenue à partir de la séquence de  $NDVI$  est très proche de celle obtenue à partir de la séquence de fraction de couvert. L'image 12.3 (c) présente, en blanc, les pixels mal labellisés par rapport à la labellisation des images de fraction de couvert **HR**. Les cartes d'erreurs présentées figures 12.2 (c) et 12.3 (c) sont très similaires, ce qui montre que l'utilisation de données de  $NDVI$  ou de frac-

---

<sup>1</sup>Normalized Difference Vegetation Index :  $(TC3 - TC2)/(TC3 + TC2)$  où  $TC2$  et  $TC3$  correspondent aux canaux 2 et 3 de **SPOT/HRVIR**, cf. figure 11.2.

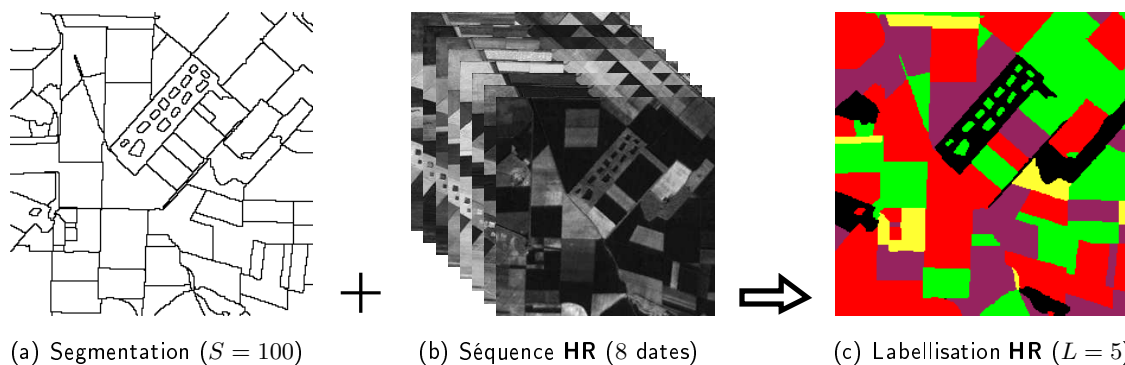


FIG. 12.1.: A partir de la segmentation en 100 régions (a) (obtenue par le module *segct* du logiciel Megawave 2 pour deux images **HR** de fraction de couvert) et de la séquence d'images de fraction de couvert (b) de même résolution, l'algorithme non-supervisé permet d'obtenir la labellisation (c) pour 5 labels (labellisation de référence).

tion de couvert fournit des résultats équivalents (dans ce cas). De plus, le faible taux d'erreurs de labellisation montre que la méthode permet, à partir d'images **BR**, d'obtenir des résultats très proches de ce qu'on obtiendrait à partir d'images **HR**.

Au-delà du débat thématique engendré par le choix d'un indice capable de discriminer au mieux les différents types de végétation, la comparaison des résultats obtenus à partir de fraction de couvert et de NDVI permet de montrer une certaine stabilité de la méthode de classification vis à vis de la mesure choisie. A noter qu'une augmentation du nombre de dates dans la séquence ou un meilleur choix des dates permettrait peut-être une amélioration des résultats.

Une autre sortie de l'algorithme non-supervisé est l'estimation des moyennes correspondant à chaque label associée à la solution. La figure 12.4 présente les moyennes estimées à partir des séquences **HR** et **BR** de fraction de couvert. Remarquons que les estimations réalisées à partir de données **BR** sont très proches de celles réalisées à partir de données **HR**. De plus, ce graphe met en évidence la nécessité d'employer des séquences d'images à travers le fait que, à une date donnée, plusieurs types d'occupation du sol soient caractérisés par un signal très proche.

A partir des relevés de terrain réalisés pour l'année agricole 2000 – 2001 (*cf.* figure 11.1), F. Oro a établi la classification des types d'occupation du sol présentée figure 12.5. Cette classification peut donc être considérée comme vérité de terrain. Douze types d'occupation du sol ont ainsi été recensés parmi lesquels dix sont représentés dans l'extrait que nous avons étudié. En mode non-supervisé, l'algorithme permet de fournir une classification **HR** à partir d'une segmentation **HR** et d'une séquence d'images **BR** de la même scène pour un nombre de labels donné. La figure 12.6 (images (c) à (l)) représente les labellisations obtenues pour 4 à 13 labels à partir de la segmentation (a) et de la séquence **BR** (b).

Remarquons que la labellisation (d) contenant 5 labels est très proche de la vérité terrain 12.5, où certains types d'occupation ne seraient pas discriminés. En effet, si la forêt, le blé, le

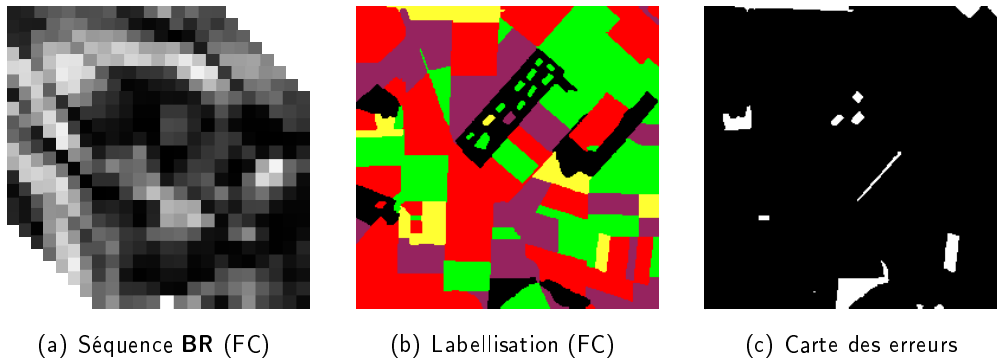


FIG. 12.2.: Labellisation obtenue pour 5 labels avec l'approche non-supervisée à partir de la segmentation (figure 12.1 (a)) et de la séquence **BR** de 8 images (a) de fraction de couvert (facteur de résolution  $16 \times 16$ ). La carte des erreurs (c) montre (en blanc) environ 3.4% d'erreurs de labellisation (pourcentage de pixels mal labellisés) par rapport à la labellisation de référence (figure 12.1 (c)).

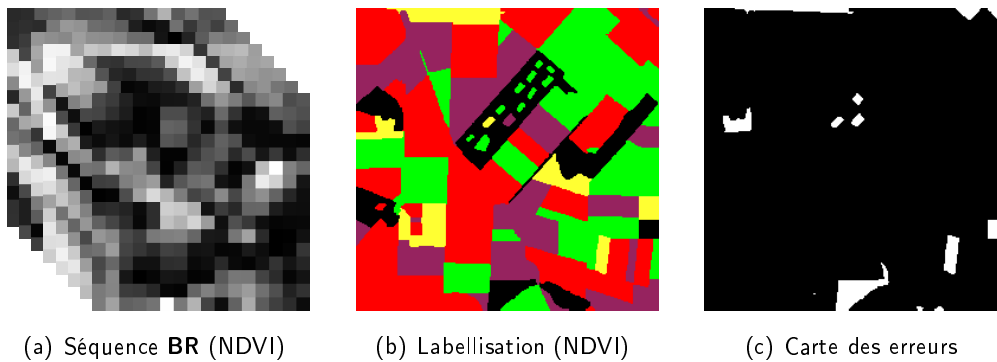


FIG. 12.3.: Labellisation obtenue pour 5 labels avec l'approche non-supervisée à partir de la segmentation (figure 12.1 (a)) et de la séquence **BR** de 8 images (b) de NDVI (facteur de résolution  $16 \times 16$ ). La carte des erreurs (c) montre (en blanc) 3.6% d'erreurs de labellisation (pourcentage de pixels mal labellisés) par rapport à la labellisation de référence (figure 12.1 (c)).



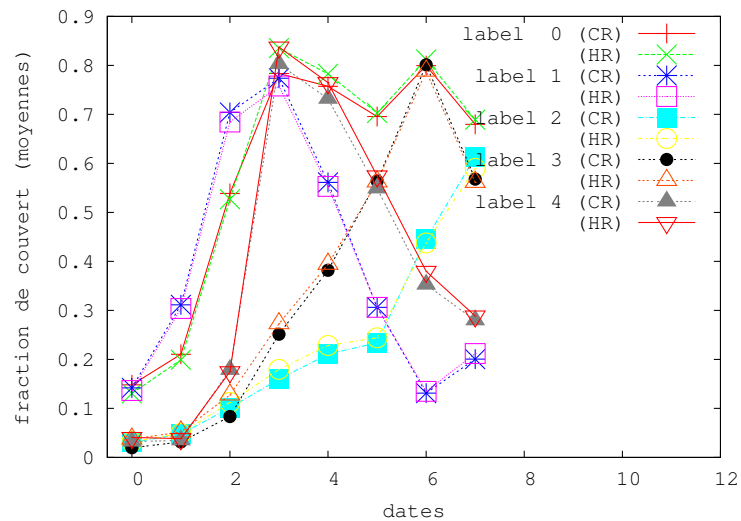


FIG. 12.4.: Evolution temporelle des moyennes de fraction de couvert estimées par l'algorithme non-supervisé à partir de la séquence **HR** et de la séquence **BR** (moyennée par blocs de taille  $16^2$ ). Les estimations réalisées à partir de données **BR** sont très proches de celles réalisées à partir de données **HR**.

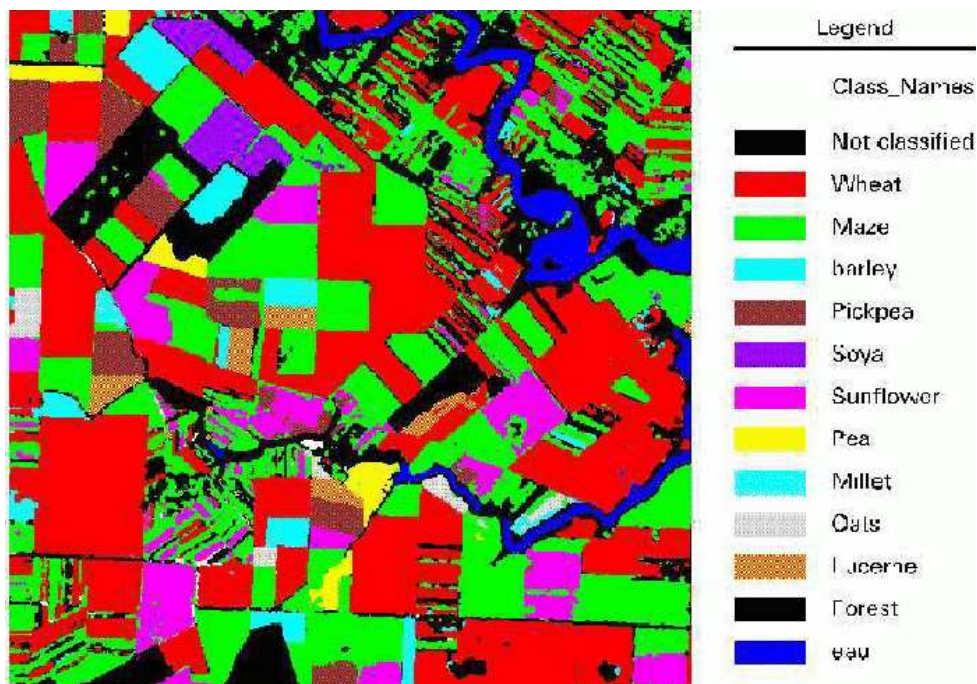


FIG. 12.5.: Classification réalisée à partir de mesures sur le terrain (F. Oro, INRA).

maïs et le pois semblent bien labellisés dans 12.6 (d), l'orge est confondu avec le blé, le soja est confondu avec le maïs, l'avoine est confondu avec le pois et le pois chiche, le tournesol et le millet sont confondus. En revanche, la luzerne est mal labellisée : elle est alternativement confondue avec la forêt, le blé ou le pois chiche selon les autres types de culture représentés dans le pixel **BR** (elle est distinguée des autres cultures à partir de 10 labels, labellisation (i)). Cette labellisation semble donc bien correspondre à la *vérité de terrain* (figure 12.5) dans la mesure où, même si chaque label correspond parfois à plusieurs types d'occupation du sol, les régions sont bien labellisées (à part la luzerne). La labellisation réalisée pour plus de labels permet de discerner davantage de types d'occupation du sol, mais certaines régions ont alors tendance à être mal labellisées.

De plus, la qualité de la segmentation a un impact direct sur celle de la labellisation puisque des parcelles non-segmentées vont nécessairement se voir attribuer le même label alors que leur type d'occupation du sol peut être différent. Dans la segmentation 12.6 (a), quelques parcelles ne sont pas séparées et, par conséquent, ne peuvent pas être labellisées différemment. Pour éviter ce type d'erreurs, il vaut mieux utiliser une sur-segmentation de la scène qu'une sous-segmentation. Cependant, le temps de calcul nécessaire à la labellisation est d'autant plus long que le nombre de régions contenues dans la segmentation est grand.

Si les erreurs de labellisation observées sur certaines parcelles (d'après la vérité de terrain figure 12.5) peuvent provenir du fait que la méthode de segmentation n'a pas séparé les régions de manière suffisamment fine, elles peuvent aussi résulter du manque de séparabilité du signal radiométrique mesuré pour les deux types d'occupation. Dans l'objectif de discerner ces deux types d'erreurs, nous réalisons les mêmes expériences à partir d'une segmentation de la même scène plus fine que celle utilisée figure 12.6 (a). La segmentation 12.7 (a) correspond à 200 régions alors que la figure 12.6 (a) en représentait 100. Certaines parcelles qui n'étaient pas séparées dans la segmentation 12.6 (a) le sont dans la segmentation 12.7 (a) mais la plupart des régions supplémentaires sont vraisemblablement situées sur les bords des régions (donc peu visibles).

Les labellisations obtenues à partir de cette segmentation pour 4 à 13 labels (figure 12.7) sont très proches de celles obtenues à partir de la segmentation en 100 régions (figure 12.6 (a)), en particulier pour 4 et 5 labels (images (c) et (d)) où elles sont quasiment identiques. De plus, remarquons que les parcelles mal labellisées qui étaient regroupées au sein d'une même région dans la figure 12.6 restent mal labellisées alors qu'elles sont bien séparées par la segmentation 12.7 (a). Ces erreurs ne proviennent donc pas uniquement de la segmentation utilisée, mais aussi du manque de séparabilité du signal radiométrique exploité ou d'une limite de la méthode.

Si les résultats présentés figures 12.6 et 12.7 semblent stables et montrent des performances satisfaisantes, ils mettent aussi en évidence la difficulté engendrée par le choix du nombre de classes pour la labellisation. Nous discuterons de ce point dans la section 12.3.

Par ailleurs, le choix des dates d'acquisition exploitées joue un rôle crucial pour la qualité de la labellisation : il est déterminant pour la séparabilité des différentes classes. Ce choix a été fait de manière arbitraire pour ces expériences, guidé typiquement par les périodes de croissance, de floraison et de sénescence : 14 novembre, 14 mars (sénescence), 12 avril, 24 et 31 mai (crois-

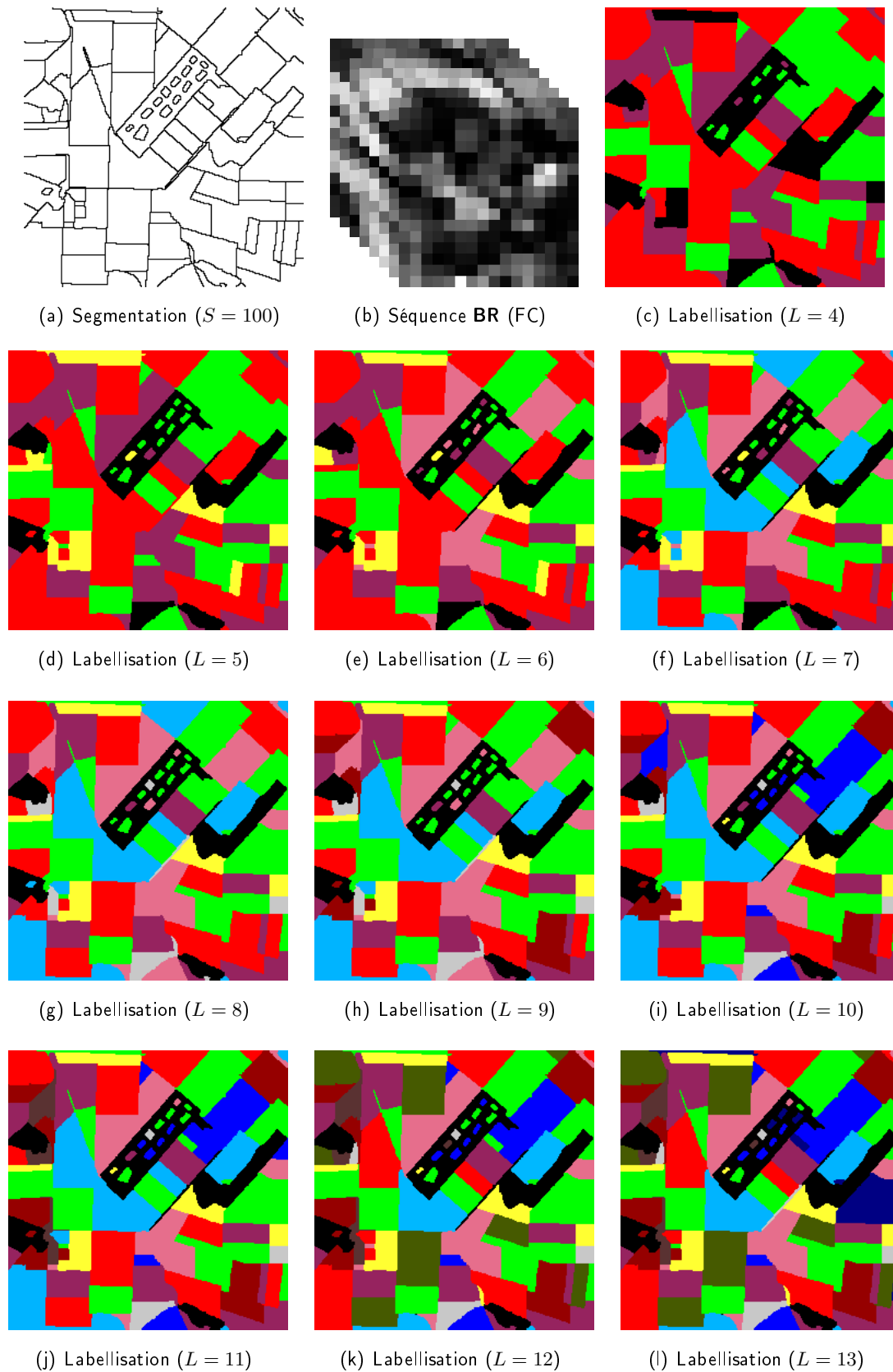


FIG. 12.6.: Labellisations obtenues en mode non-supervisé à partir de la segmentation (a) contenant 100 régions et de la séquence **BR** de fraction de couvert (b) pour 4 à 13 labels (images (c) à (l)).

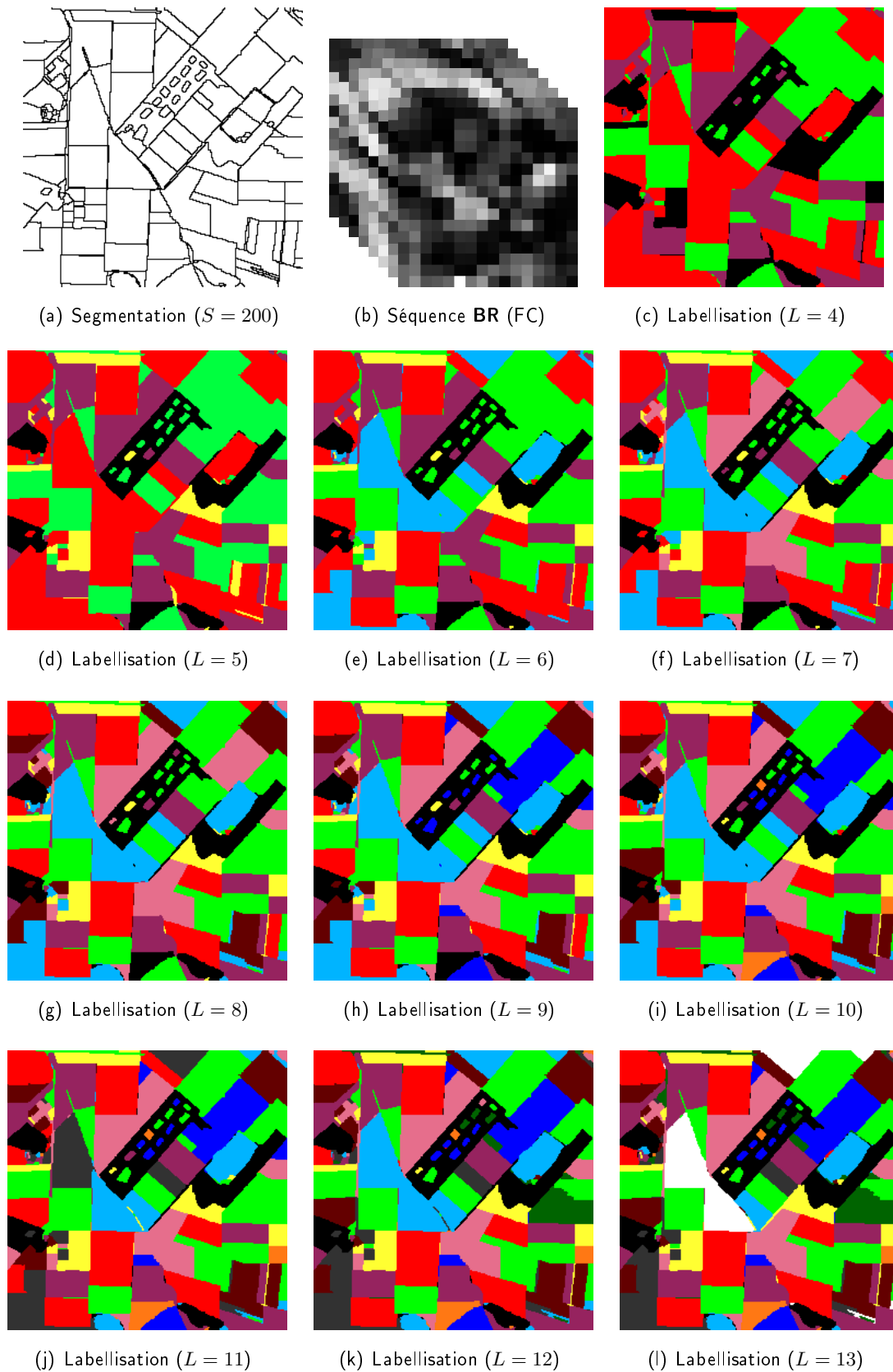


FIG. 12.7.: Labellisations obtenues en mode non-supervisé à partir de la segmentation (a) contenant 200 régions et de la séquence **BR** de fraction de couvert (b) pour 4 à 13 labels (images (c) à (l)).

sance), 13 et 24 juin, 31 juillet (floraison). Il serait intéressant d'étudier l'effet de ce choix sur les performances et d'explorer le potentiel d'une mesure telle que celle présentée dans la section 5.4 pour choisir les dates les plus pertinentes à exploiter pour la labellisation.

Dans la section 12.2, nous exploitons le critère *a contrario* défini dans le cadre de la détection de changements pour valider une labellisation à partir d'une séquence d'images **BR**.

## 12.2. Validation d'une carte d'occupation des sols

La méthode de détection de changements présentée dans la partie III peut être utilisée dans le but de valider une classification. En effet, cette méthode repose sur la définition d'une mesure de *cohérence* entre une classification et une séquence d'images **BR** et l'algorithme recherche le sous-domaine **BR** le plus cohérent avec une classification donnée. Si tout le domaine est détecté comme cohérent, nous considérons que la classification est valide.

Cette méthode de détection de changements est indépendante de la méthode utilisée pour la classification. Dans la figure 12.8 (a), nous présentons une image **HR** de fraction de couvert de taille  $1200 \times 1200$  dérivée d'une image **SPOT/HRVIR** du 11 novembre 2000 de la base de données ADAM. L'image 12.8 (b) représente la classification de l'image 12.8 (a) réalisée par méthode des  $K$ -moyennes pour 5 classes (par le logiciel de traitement d'images ENVI). Dans un contexte monotemporel, la méthode de détection de changements a été appliquée pour valider la classification 12.8 (b) à partir d'images **BR** de résolution  $5 \times 5$  à  $50 \times 50$  (images **BR** obtenues par moyennage par blocs de l'image 12.8 (a)). La figure 12.9 représente, en rouge, les pixels **BR** détectés comme non-cohérents avec la classification 12.8 (b). Si la méthode a tendance à sur-détecter les pixels non-cohérents avec une classification donnée, elle semble robuste aux changements d'échelle puisqu'elle détecte sensiblement les mêmes pixels à partir d'une image dégradée d'un facteur  $5 \times 5$  ou  $50 \times 50$ . De plus, le fait que les mêmes pixels soient détectés (figures 12.9 (a) à (f)) révèle une certaine stabilité de la détection vis à vis des pixels qui posent problème. La figure 12.9 (g) trace l'évolution du *NFA* (à l'échelle log) en fonction du facteur de résolution. Conformément à nos attentes, il est d'autant plus faible que le rapport de résolution est faible, *i.e.* la cohérence mesurée entre la classification et l'image est d'autant plus significative que le rapport de résolution est faible. En terme de pourcentage du domaine validé, la figure 12.9 (h) illustre la robustesse de la méthode vis à vis du rapport de résolution, confirmée par les images (a) à (d) pour la localisation des pixels invalidés.

Dans un contexte multi-temporel, reprenons les données et résultats de classification présentés figure 12.6. La séquence d'images est détectée comme cohérente sur 100% du domaine avec les classifications contenant de 2 à 15 labels. Plus précisément, la figure 12.10 présente les résultats obtenus sur ce jeu de données. Dans un premier temps, nous nous intéressons uniquement aux résultats de la colonne de gauche (version non-normalisée), une autre version de l'algorithme, dite normalisée, sera introduite par la suite. La figure 12.10 (a) représente l'évolution du résidu quadratique cumulé en fonction du nombre de labels. Les courbes rouge, verte et bleue corres-

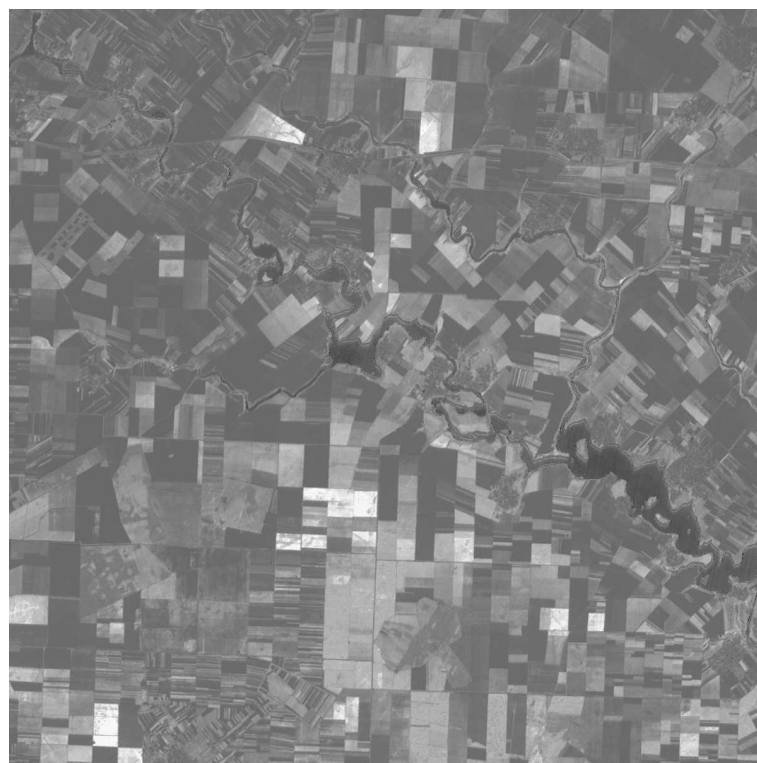
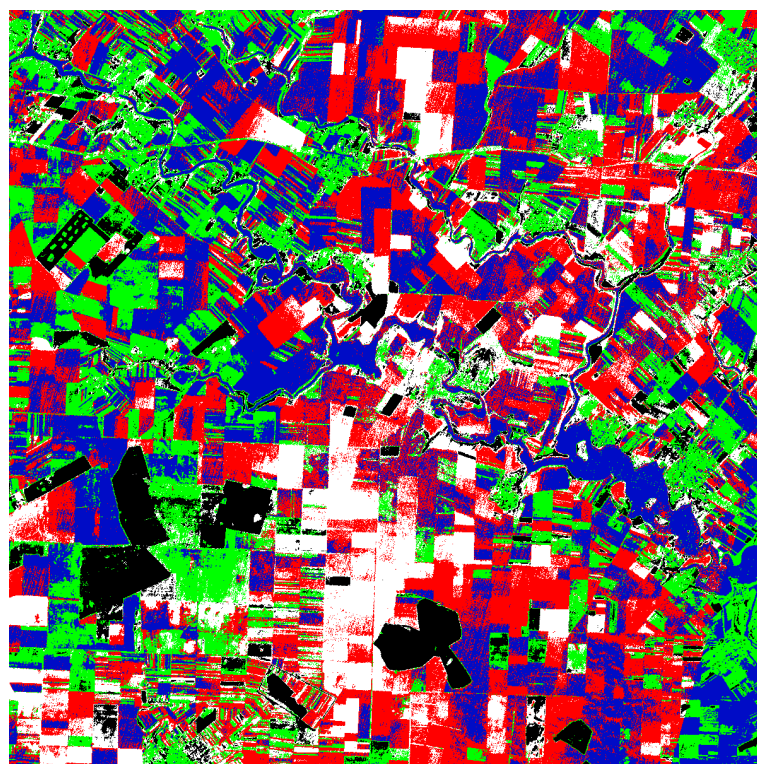
(a) Image **HR** du 11/11/00(b) Classification ( $L = 5$ ) du 11/11/00

FIG. 12.8.: Image **HR** de fraction de couvert du 14 novembre 2000 (image (a)) et classification (image (b)) réalisée par la méthode des K-moyennes pour 5 labels, à partir de l'image (a).

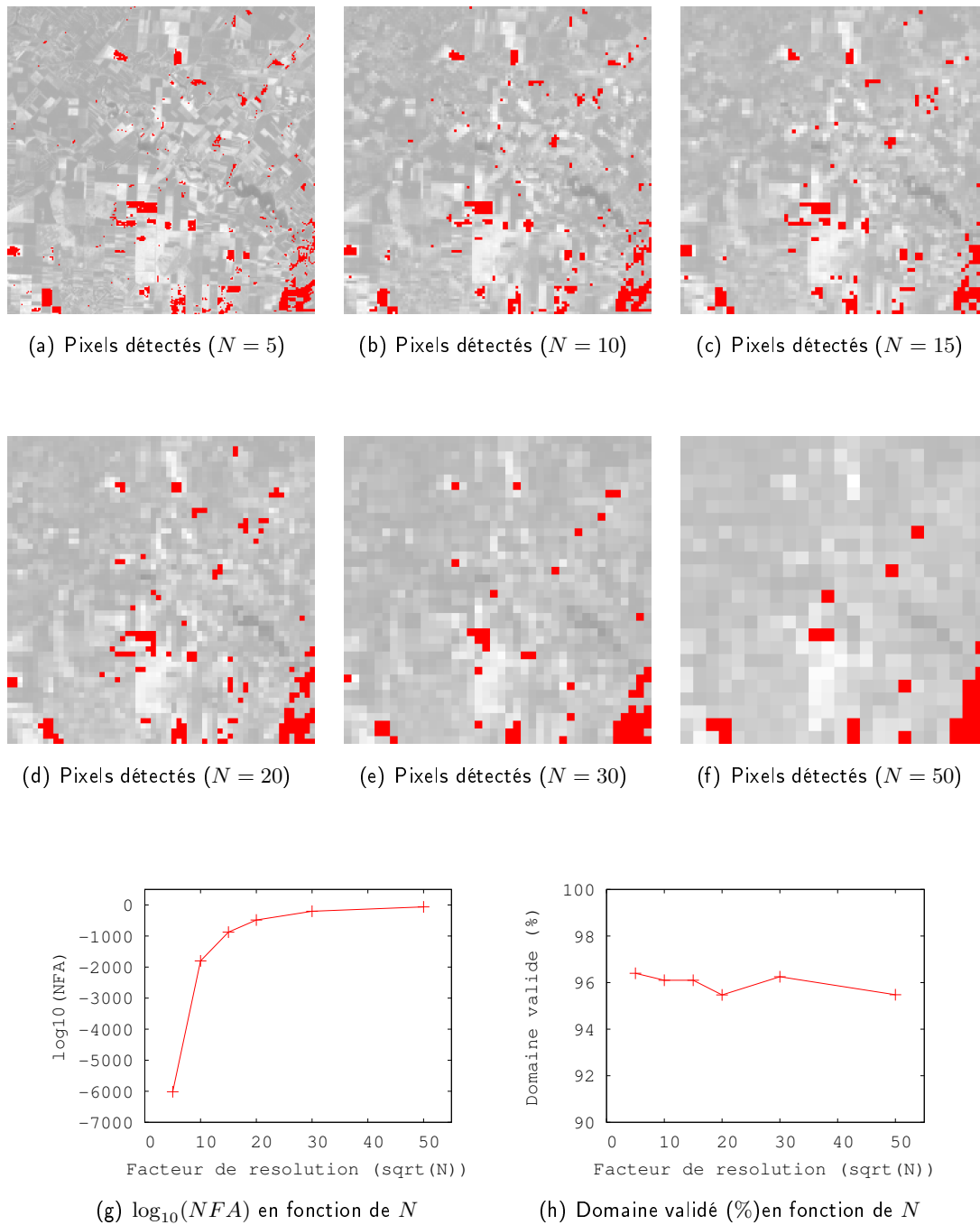


FIG. 12.9.: Validation de la classification 12.8 (b) par l'image 12.8 (a) dégradée d'un facteur  $N = 5$  à  $N = 50$  (figures (a) à (f)) avec la méthode de détection *a contrario*. Les pixels détectés comme non-cohérents avec la classification sont présentés en rouge, en surimpression sur l'image de fraction de couvert dégradée. L'évolution de la valeur du  $NFA$  (en échelle  $\log_{10}$ ) obtenue et le pourcentage du domaine validé sont tracés en fonction du facteur de résolution  $N$  (figures (g) et (h)). Le domaine est bien validé par la méthode pour 95.5 à 96.6% des pixels de l'image. Remarquons que le domaine validé est stable en fonction du rapport de résolution  $N$ .



pondent, respectivement, aux résultats obtenus avec 50 000, 500 000 et 5 000 000 itérations. Ces trois courbes, décroissantes, sont quasiment superposées pour les labellisations contenant de 2 à 10 labels, montrant la convergence de l'algorithme dès 50 000 itérations. Au-delà, le fait que le résidu augmente révèle que l'algorithme n'a pas atteint la convergence. En effet, la fonction qui à un nombre de labels  $L$  associe le résidu minimal théorique pour  $L$  labels est décroissante donc son augmentation montre que le minimum n'a pas été trouvé par l'algorithme. Les figures 12.10 (c) et (e) tracent les valeurs de  $NFA$  et le pourcentage du domaine validé correspondant. Avec un  $NFA$  minimum obtenu pour 7 à 10 labels, les labellisations les plus cohérentes avec la séquence d'images sont celles réalisées pour 7 à 10 labels. Cependant, au-delà de 10 labels, l'algorithme n'a pas convergé donc nous ne pouvons pas conclure sur la pertinence des labellisations suivantes.

Rappelons que l'algorithme adopté repose sur une stratégie de type RANSAC, *i.e.* sur le tirage aléatoire de  $L$  pixels dans chaque image de la séquence pour estimer les moyennes des classes. La sélection de ces  $L \times T$  pixels permet d'explorer tous les sous-systèmes que l'on peut former à partir de ces données. Cependant, plus le champ que l'on souhaite explorer est vaste, plus le nombre d'itérations nécessaire pour accéder à la convergence est grand.

Une alternative est de tirer aléatoirement  $L$  pixels et de considérer ces mêmes coordonnées pour chaque image de la séquence (tirage de vecteurs de pixels). Cette approche réduit considérablement le nombre de sous-systèmes explorés mais il permet certainement d'accéder à la convergence en un nombre moindre d'itérations. Par la suite, cette variante de l'algorithme sera qualifiée de vectorielle, par opposition à la version initiale non-vectorielle. La figure 12.11 présente les résultats obtenus dans le même contexte que ceux présentés figure 12.10. Remarquons que la convergence de l'algorithme semble avoir été atteinte pour les labellisations contenant de 2 à 14 labels. Cependant, cette stratégie présente des inconvénients dans le cas de données manquantes où plus le nombre de dates utilisées est grand, plus le nombre de pixels valides à toutes les dates est faible et l'algorithme vectoriel risque alors d'introduire un biais dans l'estimation des moyennes.

En dehors des questions de convergence de l'algorithme, la variance du modèle naïf est aussi un argument du modèle qui a un effet sur les résultats. Une première approche est de prendre la variance du modèle naïf égale à la variance empirique de la séquence d'images. Cependant, dans le cas où les variances des images de la séquence sont très différentes, ce modèle naïf autorise la détection de pixels de bruit. Ainsi, une telle approche peut conduire à la détection à tort de pixels simplement bruités, comme à la validation de labellisations pourtant très différentes. Par exemple, des expériences complémentaires ont montré qu'avec ce modèle naïf, des labellisations simulées par tirage aléatoire d'un label pour chaque région sont parfois validées pour tout le domaine spatio-temporel d'une séquence d'images pseudo-réelles. Une alternative pourrait être de prendre en compte les variances de chaque date dans le modèle naïf. Il faudrait, pour cela, calculer le nombre de fausses alarmes associé à ce modèle naïf, la loi de la somme de carrés de variables aléatoires gaussiennes de variances différentes n'étant pas de type  $\chi^2$ .

Pour éviter la validation de données inconsistantes, nous utilisons une version normalisée de l'algorithme : chaque image est normalisée par sa variance et la variance du modèle naïf est alors



fixée égale à 1. Les résultats obtenus avec la version normalisée sont présentés figures 12.10 (b), (d) et (f). Contrairement au *NFA* non-normalisé (cf. paragraphe précédent), le *NFA* normalisé semble évoluer par paliers en fonction du nombre de labels. En analysant visuellement les labellisations présentées figures 12.6 (a) à (j), remarquons que les changements de labels qui apparaissent entre les labellisations contenant de 5 et 6 labels correspondent majoritairement à une scission de classes (les régions labellisées en violet figure (b) sont labellisées en violet ou en rose figure (c)). De même, la prise en compte d'un label supplémentaire conduit principalement à la scission de classes dans les labellisations contenant 6, 8 et 11 labels. En revanche, les changements de label qui apparaissent sur les autres labellisations ne correspondent pas à des scissions de classes. Il semble qu'à chaque scission de classe, la valeur du *NFA* décroît assez nettement alors qu'elle est constante ou légèrement croissante dans les autres cas. Cette observation nécessiterait cependant d'être confirmée par davantage d'expériences.

De manière générale, nous nous attendons à ce que la méthode rejette tout le domaine lorsque la labellisation et la séquence d'images donnent des informations contradictoires ou lorsque la labellisation donne une information que l'on ne retrouve pas dans la séquence d'images. Remarquons que, contrairement à la version non-normalisée, une labellisation ne contenant qu'une seule classe (image constante) est alors entièrement rejetée par le modèle. De plus, pour vérifier que des labellisations quelconques ne sont pas validées par la méthode, nous avons simulé vingt labellisations par tirage aléatoire d'un label pour chaque région (à partir de la segmentation 12.6 (a)). L'algorithme appliqué à ces labellisations aléatoires et à la séquence de 8 images de fraction de couvert a conduit systématiquement au rejet de l'intégralité du domaine. Cette version de l'algorithme semble donc valider les labellisations lorsqu'elles sont correctes et les rejeter dans le cas contraire. Un autre avantage de la normalisation est qu'elle garantit de ne rien détecter dans une séquence d'images de bruit du type  $(U_1, \dots, U_T)$  où, pour tout  $t \leq T$  et pour tout  $x \in \mathcal{D}_{BR}$ , les variables aléatoires  $U_t(x) \sim \mathcal{N}(m_t, \sigma_t^2)$  sont indépendantes, contrairement à la version non-normalisée. Cependant, un inconvénient de la normalisation des images est qu'elle risque d'amplifier artificiellement le bruit des images de faible dynamique, alors que cette dernière résulte de l'absence d'information pertinente. Dans le cadre de ces expériences, les dates d'acquisition ont été choisies aux instants les plus discriminants pour les différents types de culture (sénescence, croissance, floraison, *i.e.* aux mois de novembre, mars, avril, mai, juin, juillet). L'information contenue dans chacune de ces images est donc *a priori* pertinente et les effets de l'amplification du bruit sont alors minimes.

Avec la version vectorielle non-normalisée, les labellisations contenant de 1 à 17 labels sont entièrement validées par l'algorithme, et celles contenant 18 à 20 labels sont validées à hauteur de 96 à 98% du domaine. Les figures 12.11 (c) et (d) révèlent un *NFA* minimum pour 12 labels, *i.e.* la labellisation la plus cohérente avec la séquence d'image est celle contenant 12 labels. Cette valeur minimale semble réellement significative dans la mesure où la convergence semble atteinte tant que  $L \leq 14$  (les résidus continuent de décroître jusqu'à  $L = 14$  et le nombre d'itérations ne semble pas avoir d'effet). Ce résultat ne peut donc pas être considéré comme contradictoire avec celui observé figure 12.10 (c) et (d) dans la mesure où la convergence de l'algorithme n'était pas

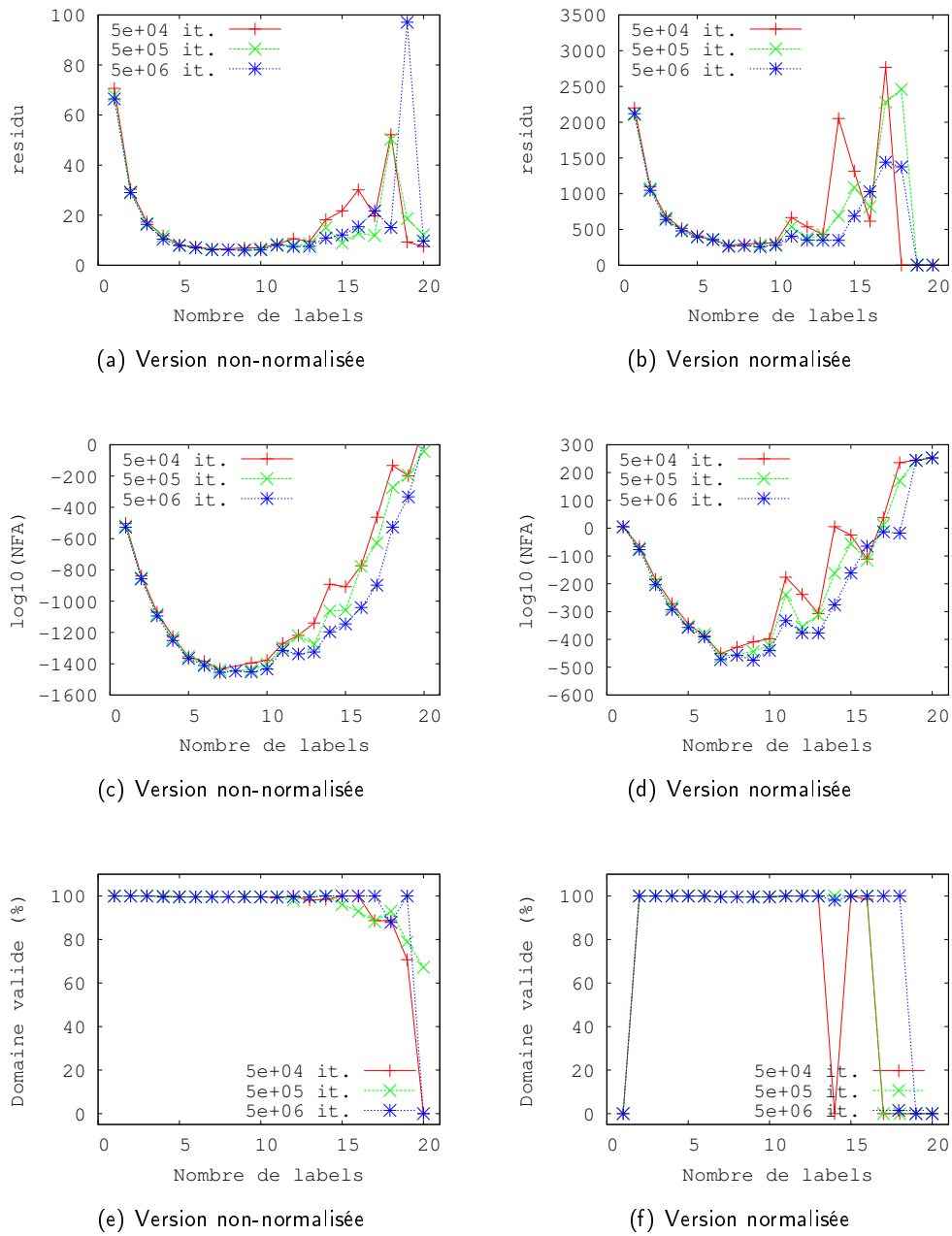


FIG. 12.10.: Résultats obtenus avec la version non-vectorielle, à partir de la séquence de 8 images et des labellisations contenant de 1 à 20 labels (figure 12.6). Les valeurs du résidu cumulé, du *NFA* (en échelle log) et du pourcentage du domaine validé sont respectivement tracées en fonction du nombre de labels contenus dans la labellisation. Chaque courbe représente les résultats obtenus pour 50 000, 500 000 et 5 000 000 itérations. Les résultats obtenus avec la version non-normalisée (modèle naïf :  $\sigma^2 =$  variance de la séquence) sont présentés à gauche et avec la version normalisée (images normalisées et  $\sigma^2 = 1$ ) à droite. Les deux versions valident les labellisations contenant de 2 à 20 labels tant que le nombre d'itérations est suffisant (au-delà de 10 labels, l'augmentation des résidus montre l'absence de convergence). La labellisation pour un seul label est rejetée, par construction, par la version normalisée.

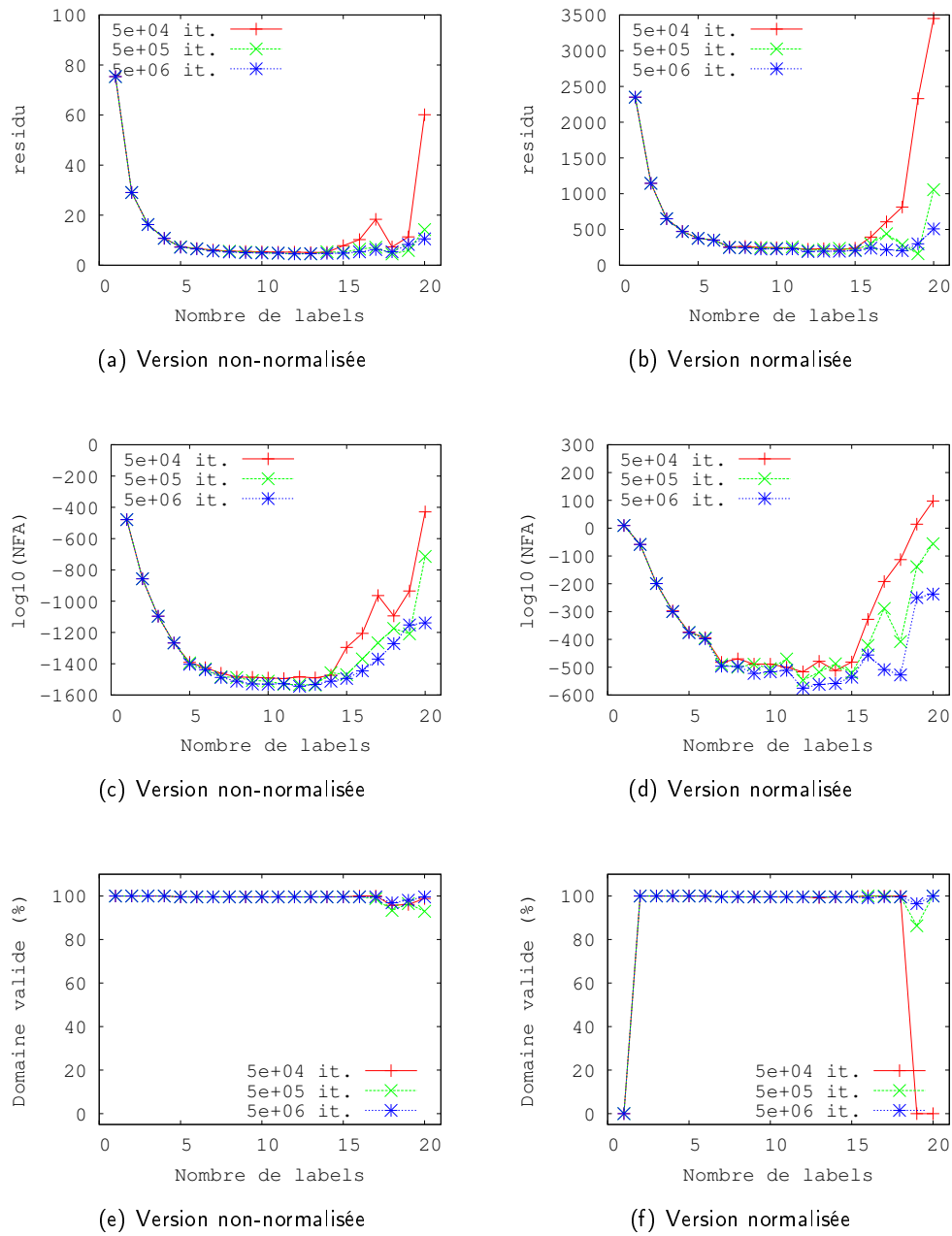


FIG. 12.11.: Résultats obtenus avec la version vectorielle, à partir de la séquence de huit images et des labellisations contenant de 1 à 20 labels (figure 12.6). Les valeurs du résidu cumulé, du *NFA* (en échelle log) et du pourcentage du domaine validé sont respectivement tracées en fonction du nombre de labels contenus dans la labellisation. Chaque courbe représente les résultats obtenus en fonction du nombre d'itérations considéré, suggérant que la convergence a été atteinte pour les labellisations contenant moins de 14 labels. Les résultats obtenus avec la version normalisée sont présentés à gauche et avec la version non-normalisée à droite. Les deux versions permettent de valider les labellisations contenant de 2 à 20 labels et marquent un minimum pour 12 labels. En revanche, seule la version normalisée permet, par construction, de rejeter la labellisation ne contenant qu'un seul label.

atteinte pour les labellisations contenant plus de 10 labels.

Compte-tenu du temps de convergence et des résultats obtenus avec la version non-normalisée dans le cas des labellisations aléatoires, nous considérerons par la suite uniquement la version vectorielle normalisée de l'algorithme. La figure 12.12 présente l'évolution des valeurs du *NFA* normalisé en fonction du nombre de labels, pour une séquence d'images donnée. Nous considérons des séquences contenant de 1 à 8 images, prises dans l'ordre des dates d'acquisitions. Chaque courbe correspond aux valeurs obtenues pour une séquence d'images : pour  $T = 1$ , seule la première date (11 novembre 2000) est considérée; pour  $T = 2$ , les deux premières dates sont considérées, etc. Remarquons que l'utilisation d'une seule date de la séquence ne permet pas de valider les labellisations (qu'il s'agisse de la première date ou des dates suivantes). Quel que soit le nombre de dates considéré, la valeur minimale du *NFA* a été obtenue pour 12 labels, désignant ainsi la labellisation contenant 12 labels comme la plus cohérente (au sens du *NFA*) avec cette séquence d'images.

Dans la section 12.3, nous discutons du problème du choix du nombre de classes *a posteriori*. Nous proposons un critère *a contrario* capable d'évaluer la *pertinence* d'une labellisation réalisée à partir d'une segmentation contenant un nombre de régions quelconque.

## 12.3. Choix du nombre de classes

La méthode de labellisation que nous avons proposée, comme la plupart des méthodes de classification (probabilistes ou non), requiert que l'utilisateur fixe au préalable le nombre de classes recherchées. Or, en pratique, on n'a généralement qu'une idée très approximative du nombre de classes réellement présentes et, même lorsque qu'il est connu, le nombre réel de classes ne correspond pas nécessairement au nombre de classes distinguables à partir des données observées. Un outil capable de guider l'utilisateur dans le choix du nombre de classes serait alors très appréciable.

Dans la littérature, différentes méthodes ont été proposées pour tenter de résoudre ce problème mais le choix du nombre de classes est, encore aujourd'hui, un problème non résolu. Il s'agit d'un problème difficile et très dépendant de l'application considérée à travers le niveau de détails recherché pour la classification. Certains algorithmes, complètement automatiques, tels que l'algorithme SEM (Stochastic Estimation Maximization, cf. [Celeux et Diebolt, 1986]) proposent de déterminer les composants d'un mélange de lois de probabilité ainsi que le nombre de ces composants. D'autres approches, plus centrées sur l'estimation du nombre de classes, sont fondées sur des divisions successives de l'échantillon de données (cf. [O'Sullivan, 1993]) et sur l'observation de l'évolution de la distance intra-classes en fonction du nombre de classes. La distance intra-classes décroît lorsque le nombre de classes augmente, mais cette décroissance est nettement plus faible au-delà du *bon* nombre de classes. La détection de l'instant à partir duquel cette décroissance est "nettement" plus faible se fait généralement visuellement. Dans le cas de modèles de mélange gaussiens, le choix du nombre de classes s'appuie souvent sur la détection (visuelle ou automatique) d'un coude dans la courbe du maximum de vraisemblance. Pour cela, la vraisemblance d'une suite

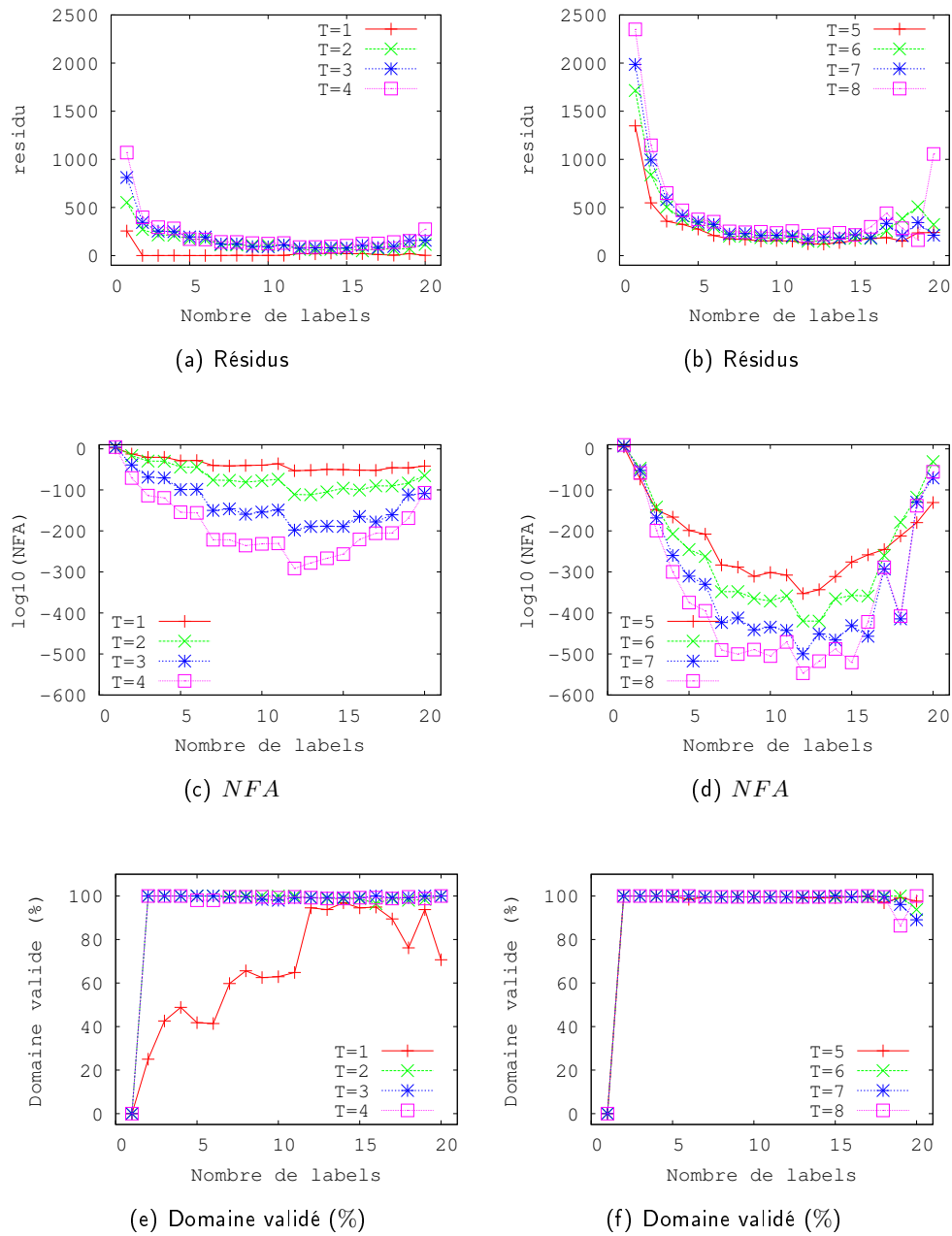


FIG. 12.12.: Résultats obtenus avec l'algorithme vectoriel normalisé pour les  $T$  premières dates de la séquence (figure 12.6 (b)) avec la labellisation (figure 12.6 (a)) réalisée pour les 8 images de la séquence. Les valeurs du résidu cumulé, du *NFA* (en échelle log) et du pourcentage du domaine valide sont respectivement tracées en fonction du nombre de labels contenu dans la labellisation. Chaque courbe représente les résultats obtenus en fonction de la séquence d'images considérée ( $T$  premières dates). La première image de la séquence ne permet pas, à elle seule, de valider les classifications. En revanche, les séquences des 2 à 8 premières images permettent de valider les classifications contenant de 2 à 15 labels (au-delà, l'augmentation des résidus montre que la convergence n'a pas été atteinte). Dans tous les cas, le *NFA* marque un minimum pour 12 labels.

de modèles (correspondant au nombre de classes, variant de 1 à  $L_{\max}$ ) est tracée en fonction de la complexité du modèle. Le principe de ces méthodes repose sur l'idée selon laquelle cette courbe forme un "coudé" lorsque le *bon* modèle est considéré. Les travaux de [Cutler et Windham, 1993] proposent de détecter automatiquement le coude de la vraisemblance au travers du critère EL (Elbow Likelihood), en introduisant un seuil dont la détermination reste très empirique. Dans le même esprit, [Biernacki, 1999] propose d'exploiter des informations supplémentaires sur les données pour détecter un seuil dans la courbe du maximum de vraisemblance et choisir, ainsi, le nombre de classes. Au-delà du fait que cette méthode nécessite des informations complémentaires, elle propose différents nombres de classes selon la précision des informations prises en compte. Globalement, les performances de ces méthodes sont variables selon les applications considérées et ne fournissent que des solutions partielles dans le sens où le choix d'un seuil (ou, dans le cas des travaux de [Biernacki, 1999], la précision des informations à intégrer) reste à déterminer.

Le nombre de fausses alarmes utilisé dans la section 12.2 pour la validation d'une classification correspond à une mesure de significativité d'une classification vis à vis d'une séquence d'images. Cette valeur a un sens de manière absolue. Elle permet, notamment, de comparer la significativité de différentes classifications (contenant différents nombres de classes) vis à vis d'une image ou séquence d'images. Dans cette section, nous explorons le potentiel du nombre de fausses alarmes pour estimer le nombre de classes.

Dans cet objectif, nous nous plaçons dans un contexte où le nombre de classes correspondant réellement au nombre de classes distinguables dans l'image. Pour cela, nous considérons un jeu de quinze labellisations, contenant de 5 à 20 labels, ainsi qu'une image simulée par tirage gaussien à partir de la labellisation contenant 10 labels puis dégradée (par moyennage par blocs de taille  $16 \times 16$ ). La figure 12.13 présente les valeurs du *NFA* obtenues pour cette image **BR** et chacune des labellisations (en échelle log) en fonction du nombre de labels contenu dans ces labellisations. Remarquons que la valeur minimale du *NFA* est obtenue pour la labellisation contenant 10 labels. Pour une image ou séquence d'images fixée, la labellisation qui contribue le plus à *expliquer* la cohérence de l'image peut donc être détectée en recherchant celle qui minimise le *NFA*.

Dans le cas des images réelles, les figures 12.10 et 12.11 montrent que l'évolution du *NFA* marque un minimum pour 12 labels, ce qui correspond au nombre de classes de la vérité de terrain présentée figure 12.5. Le fait que la courbe du *NFA* admette un minimum annonce des perspectives intéressantes pour l'estimation du nombre de classes, en se démarquant de méthodes telles que celles basées sur la courbe de vraisemblance, strictement décroissantes en fonction du nombre de labels. Cependant, ce nombre de fausses alarmes a été défini pour un nombre de classes fixé.

Dans l'objectif de déterminer le nombre de labels le plus cohérent avec une séquence d'images, nous proposons de définir un critère de pertinence d'une labellisation à  $L$  labels dans un ensemble de labellisations contenant un nombre quelconque de labels (réalisées à partir d'une segmentation à  $S$  régions) pour une séquence de  $T$  images. Ce critère est dérivé du modèle de

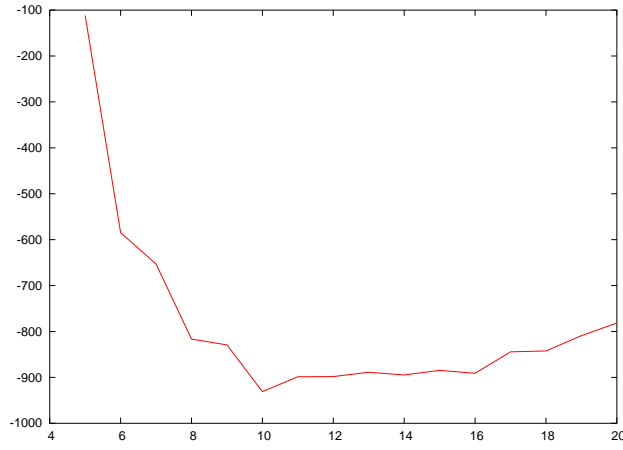


FIG. 12.13.:  $NFA$  minimal (en échelle log) obtenu pour des labellisations contenant de 5 à 20 labels et une image **BR** simulée à partir de 10 labels. Le minimum est bien obtenu pour 10 labels.

détection de changements (cf. section 7).

**Proposition 12.3.1** Soient  $L_{\max}$  le nombre maximal de labels autorisés pour la labellisation,  $E_{S,L}$  l'ensemble des labellisations créées à partir d'une segmentation en  $S$  régions avec au plus  $L$  labels et  $\lambda$  une labellisation de  $E_{S,L}$ . Etant donnée  $v$  une séquence de  $T$  images définie sur un domaine spatio-temporel  $\Omega$ , la relation

$$NFA_2(L, S, \Omega, \delta(v_\Omega), \sigma) = L_{\max} L^S \Gamma_{inc} \left( \frac{|\Omega| - L \times T}{2}, \frac{\delta^2(v_\Omega)}{2\sigma^2} \right), \quad (12.3.1)$$

où  $\delta^2(v_\Omega)$  représente le résidu quadratique cumulé sur la séquence d'images, définit un nombre de fausses alarmes sous l'hypothèse  $H_0$ .

Soit  $\varepsilon > 0$ , une labellisation de  $E_{S,L}$  est dite  $\varepsilon$ -**significative** pour une séquence d'images  $v$  si  $NFA_2(L, S, \Omega, \delta(v_\Omega), \sigma) \leq \varepsilon$ .

**Démonstration.**

D'après la proposition 7.2.4 et le théorème 7.2.5, le  $NFA_2$  définit un nombre de fausses alarmes si

$$\sum_{\lambda \in E_{S, L_{\max}}} \frac{1}{L_{\max} L^S} \leq 1.$$

Or

$$\sum_{\lambda \in E_{S, L_{\max}}} \frac{1}{L_{\max} L^S} = \frac{1}{L_{\max}} \sum_{L=1}^{L_{\max}} \sum_{\lambda \in E_{S,L} \setminus E_{S,L-1}} \frac{1}{L^S} \quad (12.3.2)$$

$$= \frac{1}{L_{\max}} \sum_{L=1}^{L_{\max}} \frac{|E_{S,L} \setminus E_{S,L-1}|}{L^S} \quad (12.3.3)$$

$$= \frac{1}{L_{\max}} \sum_{L=1}^{L_{\max}} \frac{L^S - (L-1)^S}{L^S}, \quad (12.3.4)$$

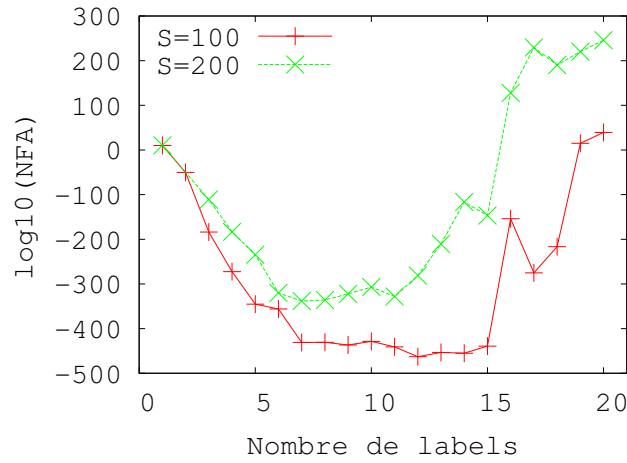


FIG. 12.14.: Evolution du  $NFA_2$  en fonction du nombre de labels. Résultats obtenus pour la séquence de huit images de fraction de couvert (figure 12.2 (a)) et les labellisations contenant 1 à 20 labels présentées figure 12.6 (en rouge) et 12.7 (en vert) (réalisées respectivement à partir de segmentations contenant 100 et 200 régions). La valeur minimale est atteinte pour 12 labels dans le cas des labellisations 12.6 ( $S = 100$ ), indiquant que la labellisation la plus pertinente pour la séquence d'images est celle réalisée avec 12 labels. Dans le cas des labellisations présentées figure 12.7 ( $S = 200$ ), un minimum est marqué pour 11 labels mais aussi pour 7 labels (valeur très proche). La pertinence des labellisations réalisées à partir de 200 régions est cependant plus faible que celle des labellisations réalisées à partir de 100 régions.

car  $L^S$  représente le nombre de labellisations contenant au plus  $L$  labels. Par conséquent,

$$\sum_{\lambda \in E_{S, L_{\max}}} \frac{1}{L_{\max} L^S} \leq 1$$

et le  $NFA_2$  définit un nombre de fausses alarmes. □

Remarquons que la variable  $L_{\max}$  n'a aucune influence sur la valeur du  $NFA_2$ , elle n'est donc pas un paramètre du modèle.

La figure 12.14 représente l'évolution du critère de pertinence ( $NFA_2$ , en échelle log) calculé pour les labellisations présentées figures 12.6 (en rouge, avec 100 régions) et figures 12.7 (en vert, avec 200 régions) en fonction du nombre de labels qu'elles contiennent. La valeur minimale pour les labellisations obtenues à partir de 100 régions est atteinte pour 12 labels, indiquant que la labellisation réalisée pour 12 labels est la plus pertinente (relativement aux autres labellisations, réalisées avec 2 à 20 labels). Le critère  $NFA_2$  calculé pour les labellisations réalisées à partir de 200 marque un minimum pour 7 et pour 11 labels (valeur du  $NFA_2$  très proche pour 7 et 11). Plus généralement, remarquons que les valeurs du  $NFA_2$  calculées pour les labellisations obtenues



à partir de 100 régions sont plus faibles que celles obtenues à partir de 200 régions, montrant que la pertinence des labellisations réalisées à partir de 100 régions est globalement supérieure aux autres. Ces résultats annoncent un nombre de classes très proche du nombre de classes recensées dans le relevé de terrain (12 classes). Cependant, davantage d'expériences sont nécessaires pour s'assurer que le minimum du  $NFA_2$  correspond à un nombre de classes pertinent pour les données considérées.

## 13. Détection de changements et mise à jour de classification

L'algorithme de détection de changements présenté dans la partie III prend en entrée une classification de référence et une séquence d'images **BR**. Il retourne le domaine **BR** le plus cohérent avec la classification de référence. La recherche de ce domaine peut donc être utilisée pour valider une classification, pour détecter les pixels candidats aux changements, ou dans un scénario de mise à jour de classification.

### 13.1. Détection de changements

Dans cette section, nous appliquons la méthode de détection *a contrario* (cf. Partie III) à deux extraits de la zone d'étude du projet ADAM. Si des relevés de terrain ont permis d'évaluer la méthode de labellisation en identifiant les types d'occupation du sol, aucune information de terrain n'est disponible sur les changements qui ont pu se produire au cours de la période d'intérêt. Sans information de ce type, nous proposons de simuler des changements de différents types pour évaluer les performances de la détection dans un contexte réaliste.

Tout d'abord, étant donnée une labellisation associée à une séquence d'images, nous proposons d'introduire des changements sur la labellisation en attribuant un nouveau label à quelques régions de la segmentation (figure 13.1 (a)) tirées aléatoirement. La figure 13.1 présente les résultats obtenus avec la méthode *a contrario* en considérant la séquence de 8 images de fraction de couvert présentée figure 12.2 (a) et la labellisation correspondante pour 10 labels (figure 13.1 (b)), modifiée en quelques régions. Dans la section 12.2, la méthode *a contrario* appliquée à la validation de la labellisation 13.1 (b) par la séquence des huit images **BR** a conduit au rejet d'un pixel du domaine **BR** (cf. figure 13.1 (c)), en tant que pixel ne contribuant pas à maximiser la cohérence de la labellisation avec la séquence d'images. Nous considérons donc ce pixel comme *invalidé* pour les expériences qui suivent. Les changements introduits dans la labellisation sont représentés en vert figures 13.1 (d), (e) et (f), en surimpression sur une image **BR** de la séquence. Sur ces mêmes images, les pixels détectés sont représentés en rose lorsqu'il s'agit de pixels *invalidés* (pixels déjà détectés avant l'introduction de changements) et en rouge sinon. Les zones de l'image qui correspondent à la fois à des zones de changements et à des pixels détectés sont représentés en jaune.

Rappelons que la méthode détecte des pixels **BR** et que les changements sont introduits

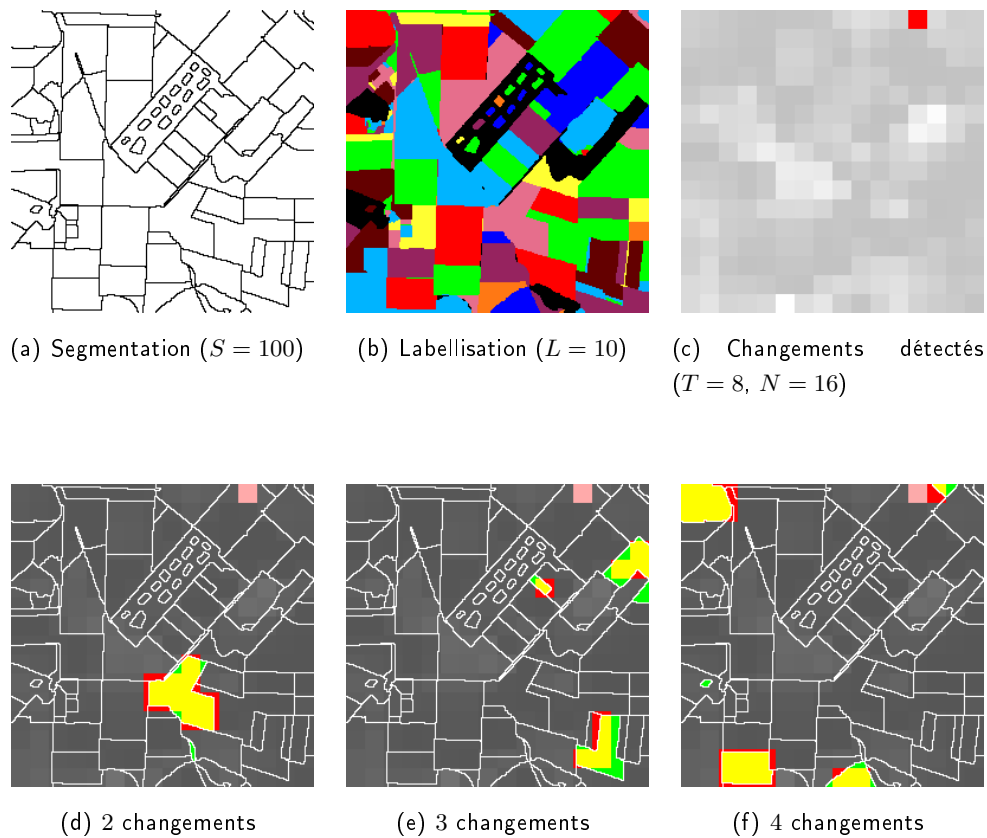


FIG. 13.1.: Résultats obtenus pour des changements introduits dans la classification par tirage aléatoire de 3, 4 ou 5 régions, et d'un nouveau label pour les régions tirées. Le label est tiré entre 1 et  $L$ . Les changements introduits dans la labellisation sont représentés en jaune lorsqu'ils sont détectés et en vert sinon. Les pixels détectés qui ne correspondent pas à un changement sont représentés en rose s'ils étaient déjà détectés avant l'introduction de changements (*cf.* (c)) et en rouge sinon. Les changements introduits sont globalement bien détectés, même lorsqu'ils concernent une faible portion d'un pixel **BR**. Dans l'exemple figure (c), un pixel de changement n'a pas été détecté (en haut à droite, en vert).

sur une labellisation **HR**. C'est la raison pour laquelle des zones rouges apparaissent souvent sur les contours des zones de changements bien détectés (jaune). La plupart des changements, mêmes lorsqu'ils sont de taille inférieure à celle du pixel **BR**, sont bien détectés. Cependant, remarquons, dans l'exemple présenté figure 13.1 (e), que des zones de changements (en vert) n'ont pas été détectées. Elles contribuent vraisemblablement à la cohérence de la labellisation, pour la mesure de cohérence introduite. Globalement, les changements introduits sont bien détectés et, à part le pixel *invalidé* dès le départ, aucun pixel n'est détecté comme changement à tort.

Dans l'objectif d'envisager un autre type de changements, nous considérons la labellisation 13.2 (a) réalisée avec une image de fraction de couvert acquise le 11 novembre 2000 et cette même image modifiée sur deux zones (représentées en blanc figure 13.2 (c)) en insérant des valeurs réelles de fraction de couvert obtenues en avril 2001. La méthode de détection appliquée à la labellisation 13.2 (a) et à l'image 13.2 (b) (comportant des changements) permet de détecter les pixels représentés en rose et en rouge figure 13.2 (c). Les pixels roses correspondent aux pixels qui étaient déjà détectés lors de la validation de la labellisation par l'image du 11 novembre. En rouge, les pixels détectés sont principalement concentrés dans les zones de changements (encadrées en blanc).

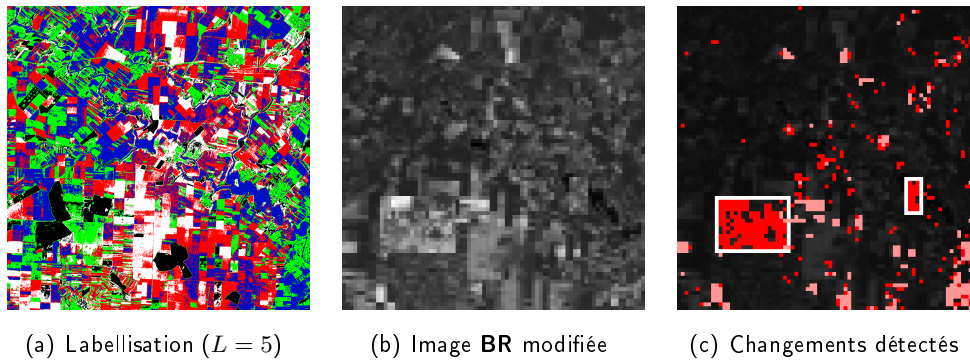


FIG. 13.2.: Détection de changements introduits sur l'image **BR** (b) correspondant à la classification (a) : les changements détectés sont représentés en rouge dans la figure (c) et la trace des changements introduits est représentée en blanc sur la même image. Les changements ont été introduits sur 5.46% des pixels de l'image **BR** et 96.2% ont été invalidés (*cf.* figure 12.9 (c)). Les pixels détectés représentent 89.3% des pixels de l'image, valeur très proche des 90.7% attendus.

## 13.2. Mise à jour de classification

Dans cette section, nous proposons d'étudier le potentiel du *NFA* dans un scénario de mise à jour de classification. Dans cet objectif, nous considérons la zone extraite présentée dans la figure 12.6 et la labellisation réalisée pour 10 labels (i) pour l'année 2000 – 2001. Dans le cas de

zones agricoles telles que celle-ci, les labellisations changent généralement d'une année à l'autre en raison de la rotation des cultures. En l'absence de données correspondant à l'année précédente, nous proposons de simuler une labellisation qui correspondrait à l'année précédente en modifiant le label de certaines parcelles. La figure 13.3(a) représente la labellisation de l'année 2000 – 2001 pour 10 labels, et la figure (b) la labellisation modifiée artificiellement en tant que labellisation de l'année 1999 – 2000. Les différences entre ces labellisations sont localisées en vert dans l'image (c). La méthode *a contrario* appliquée à la labellisation (b) et à des séquences de 1 à 8 images réparties sur l'année agricole 2000 – 2001 permet, dès la troisième image acquise, de détecter correctement les zones de changements. Les valeurs de *NFA* obtenues et le pourcentage du domaine validé sont tracés en fonction du nombre de dates utilisées pour la détection figures (d) et (e), montrant que seuls 91% (environ) des pixels du domaine sont validés comme cohérents avec la labellisation dès que les 3 premières images sont exploitées. Les figures (f) à (m) présentent, en rouge, les pixels **BR** détectés en fonction du nombre de dates considéré. A partir de la première image et des deux premières images, les pixels de changements sont largement sur-détectés. Cependant, cette non-validation d'une grande partie du domaine est davantage liée à l'inconsistance des données d'hiver, pauvres en information discriminante, qu'à la présence de changements. En effet, la majorité des parcelles est alors en sols nus alors que la labellisation distingue les différents types de culture. Notons que ce point était déjà visible sur la figure 12.12 où, sans changement, une grande partie du domaine **BR** était détectée comme incohérente avec la labellisation lorsque seule la première date était considérée. Ceci est dû, vraisemblablement, au fait que la labellisation soit trop informative par rapport à l'observation. Par ailleurs, s'agissant des dates du 14 novembre 2000 et du 13 Mars 2001, il serait prématuré de réaliser une nouvelle classification pour l'année agricole 2000 – 2001 à partir de ces seules dates. A partir des 3 premières dates, le domaine de changements est bien détecté indiquant qu'une mise à jour de la labellisation serait nécessaire.

Dans le contexte d'une analyse intra-annuelle, différents types de cas peuvent se présenter. En effet, une classification qui aurait été réalisée au cours des premiers mois d'une année agricole pourrait perdre sa validité avec l'acquisition de nouvelles images si celles-ci permettent de distinguer des types d'occupation du sol qui n'étaient pas discriminés lors de la classification initiale. Dans ce cas, les changements correspondent à des scissions de classe. D'autres types de changements peuvent survenir, tels qu'une inondation, un incendie, une infection ou une coupe de forêt. Ces derniers types de changements impliquent réellement que la classification n'est plus valable et que les acquisitions précédentes ne sont plus exploitables. En revanche, dans le cas des scissions de classe, les acquisitions précédentes peuvent encore être exploitées en complément des acquisitions suivantes pour affiner la classification. Par ailleurs, notons que dans certains cas d'application, il peut être plus judicieux d'exploiter la méthode de détection de changements dans un contexte monotemporel.

Dans la figure 13.4, nous présentons un exemple de détection de changement intra-annuel. Pour cela, nous considérons la labellisation (a) réalisée à partir des 4 premières dates de la séquence pour 12 labels et nous utilisons les 4 dates suivantes pour le suivi de la surface. Nous simulons une zone de changements (du type incendiée) sur les acquisitions 5, 6, 7 et 8 (*cf.* figures 13.4 (b) et (c))

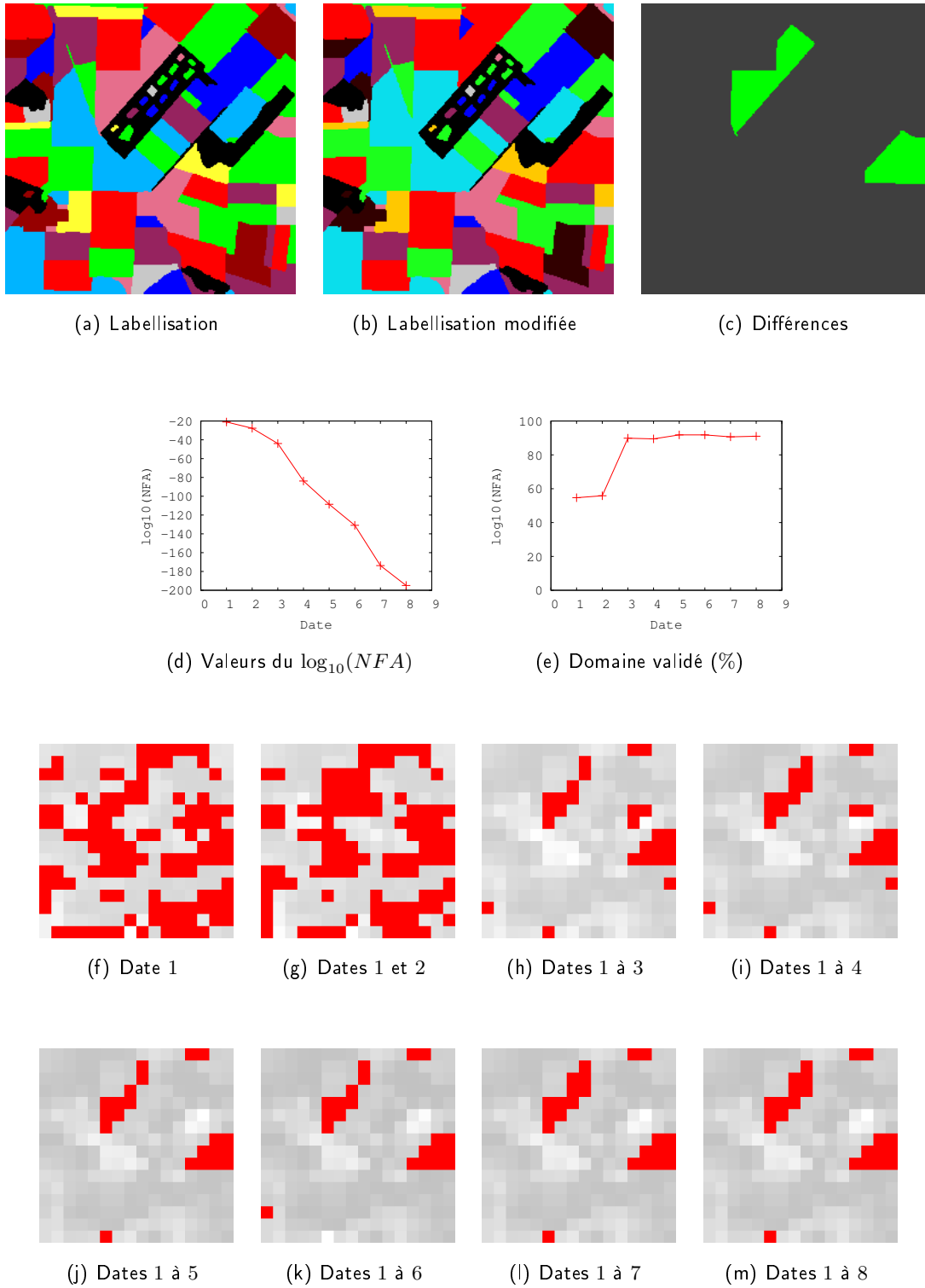


FIG. 13.3.: La labellisation (b) est simulée à partir de (a) en tant que labellisation de l'année précédente. L'image (c) présente en vert les différences entre les deux labellisations. La valeur du  $NFA$  et du pourcentage du domaine validé sont tracés en fonction du nombre de dates considéré. Le  $NFA$  permet, dès que 3 dates sont observées, de détecter les pixels qui ne sont pas cohérents avec la labellisation (en rouge, images (f) à (m)).

par tirage aléatoire gaussien respectivement de moyenne 0.02, 0.04, 0.07 et 0.1. La méthode *a contrario* appliquée, dans un contexte monotemporel, à la labellisation (a) et successivement aux images des dates 5 à 8 permet de détecter les pixels représentés en rouge figures 13.4 (h) à (k). Les changements introduits dans les images 5 et 6 sont bien détectés (lorsque la différence de niveau de gris engendrée par l'introduction de changements n'est pas trop faible). Aux dates suivantes, de nombreux changements sont détectés. A titre de comparaison, nous appliquons la méthode aux mêmes images sans changement. Les résultats présentés figures (d) à (g) montrent que les images 5 et 6 étaient globalement bien validées par la labellisation, au contraire des dates suivantes qui étaient déjà majoritairement invalidées pour la labellisation (a). Remarquons que la plupart des pixels détectés aux dates 7 et 8 à partir des images contenant des changements (images (j) et (k)) étaient déjà détectés en l'absence de changements (images (f) et (g)). En effet, dans un contexte multitemporel, la prise en compte des acquisitions précédentes peut favoriser la cohérence entre une labellisation et une séquence d'images et empêcher la détection de pixels de changements qui auraient pu être détectés en monotemporel. Pour éviter ce type d'effet, il pourrait être intéressant de considérer la détection de domaines de changements spatio-temporels en autorisant, par exemple, le rejet d'une zone spatiale pour une seule date de la séquence.

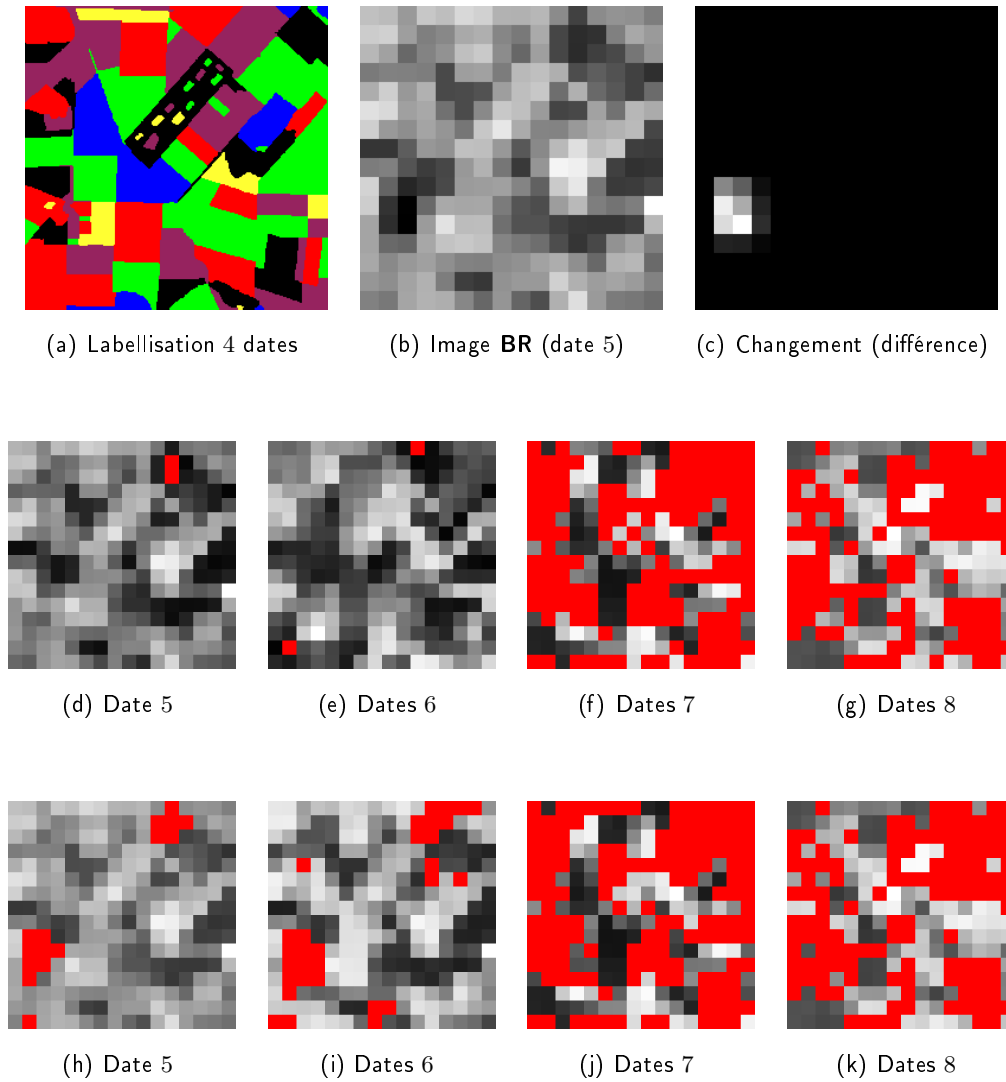


FIG. 13.4.: Simulation d'un changement sur les 4 dernières images d'une séquence de 8 acquisitions. La labellisation (a) (6 labels) correspond uniquement aux 4 premières dates. Les images (b) et (c) présentent, respectivement, l'image **BR** de la date 5 contenant un changement et l'image de différence faisant apparaître le changement introduit en niveaux de gris. Dans un contexte monotemporel, les images (d) à (g) représentent en rouge les pixels détectés en comparant la labellisation (a), successivement, aux images **BR** des dates 5, 6, 7 et 8 avant l'introduction de changements. Les images (h) à (k) représentent en rouge les pixels détectés à partir des images contenant les changements. Les pixels de changement (suffisamment marqués, *cf.* (c)) sont bien détectés à partir des dates 5 et 6. Ensuite, de nombreuses détections sont observées, que des changements aient été introduits ou non (figures (f), (g), (j) et (k)). Ces résultats semblent plutôt raisonnables puisque la labellisation considérée a été réalisée uniquement avec les 4 premières dates (d'hiver), une mise à jour est alors vraisemblablement nécessaire.





# Conclusion générale

Dans ce travail de thèse, nous nous sommes intéressés à la classification et à la détection de changements à partir de séquences d'images satellitaires pour l'analyse et le suivi des surfaces continentales. La nécessité de considérer deux résolutions spatiales (haute et basse résolutions) pour l'acquisition des données a conduit à la définition d'une problématique spécifique, avec notamment l'existence de pixels mixtes.

L'approche que nous avons proposée pour la classification repose sur une modélisation bayésienne du problème et sur le modèle linéaire de mélange pour décrire la séquence d'images **BR** en terme du mélange des différentes régions de la segmentation présentes dans chaque pixel **BR**. Alors que la problématique que nous avons considérée est liée à l'utilisation de données basse résolution, nous nous sommes intéressés à l'influence de la basse résolution sur les solutions en comparant les problèmes formulés à haute et basse résolution. Dans un contexte théorique, nous avons proposé une évaluation de la perte d'information engendrée par la basse résolution calculable à partir de l'image **HR** et du rapport de résolution considéré. Ce résultat permet, étant donné un jeu de données, de quantifier par avance la dégradation de la labellisation provoquée par le passage de la haute résolution à la basse résolution.

Par ailleurs, une analyse probabiliste nous a permis de mettre en évidence les principaux facteurs d'erreurs de labellisation. Conformément à nos attentes, l'évolution temporelle des observations moyennes caractéristiques de chaque classe joue un rôle crucial pour la fiabilité d'une labellisation. En pratique, les dates sont souvent choisies selon le stade de croissance de la végétation, ce qui ne garantit pas une labellisation optimale. L'analyse que nous proposons permet de mettre en évidence un critère pour le choix des dates (et du nombre de dates) qui devrait permettre d'optimiser la labellisation.

Nous avons adopté un algorithme de recuit simulé pour minimiser l'énergie et obtenir une labellisation à partir d'une segmentation **HR** et une séquence d'images **BR** de manière non-supervisée. Une analyse empirique des erreurs de labellisation obtenues à partir de données simulées montre que la convergence de l'algorithme (propriété délicate pour ce type de méthodes) ne semble pas poser de problème, bien qu'une loi géométrique ait été adoptée pour la décroissance de température.

D'un point de vue empirique, les expériences menées sur des données simulées ont montré une stabilité des performances de l'algorithme non-supervisé, jusqu'à un rapport de résolution de  $30 \times 30$ . Ces résultats ne permettent pas d'espérer des résultats de classification satisfaisants avec l'utilisation conjointe de données **SPOT/HRV** et **VGT** (rapport de résolution  $50 \times 50$ ). En

revanche, les résultats obtenus en exploitant des données réelles **SPOT/HRV** et **MeRIS** simulées (rapport de résolution  $15 \times 15$ ) dans le contexte de l'application à une zone agricole de la Plaine du Danube (base de données ADAM) semblent bien correspondre aux relevés de terrain disponibles.

Les expériences menées dans un cadre réaliste ont mis en évidence la difficulté du choix d'un nombre de classes recherchées. Pour tenter d'automatiser ce choix, nous avons proposé un critère de pertinence d'une labellisation représentant un nombre de classes quelconque. Les premiers résultats obtenus à partir d'un jeu de données réelles montrent que ce critère marque nettement un minimum pour un nombre de classes qui semble très proche du nombre de classes recensées sur le terrain. Une étude plus approfondie est toutefois nécessaire pour étudier la bonne correspondance entre ce minimum et le nombre de classes recherchées en réalité de manière plus générale.

D'autre part, nous avons étudié le problème de la détection de changements basée sur la comparaison d'une séquence d'images basse résolution à une classification haute résolution décrivant l'état de la surface à la date de référence. Cette formulation autorise *a priori* la détection de changements de taille sous-pixelique. De plus, la prise en compte d'une classification de référence plutôt que d'une séquence d'images permet de minimiser l'*a priori* sur l'évolution attendue.

La méthode que nous avons proposée repose sur la définition d'un critère de cohérence d'une séquence d'images **BR** et une classification **HR**, fondé sur un modèle *a contrario* qui ne nécessite pas d'information *a priori*. En effet, plutôt que de modéliser les données *a priori*, la méthode repose sur le rejet d'un modèle naïf (non structuré) par l'observation de données structurées. Une analyse théorique de ce modèle a permis de mettre en évidence un certain nombre de propriétés annonçant ses performances en fonction, notamment, du niveau de contraste de l'image ainsi que du nombre de pixels de changements de l'image.

La mise en œuvre de ce modèle nécessite l'estimation des moyennes associées à chaque classe ainsi que la recherche du sous-domaine spatio-temporel qui minimise le *NFA*. Pour cela, nous avons adopté un algorithme d'échantillonnage aléatoire (de type RANSAC) qui permet d'estimer ces moyennes avec une grande robustesse même en présence de nombreux pixels "aberrants". Alors que les performances des méthodes existantes sont généralement limitées dès que la proportion des pixels dépasse, typiquement, 20% des pixels de l'image, l'approche que nous proposons permet d'outrepasser largement ces limites, avec de bonnes performances même en présence de 70% de pixels aberrants.

Dans l'objectif d'optimiser l'algorithme, nous avons proposé plusieurs stratégies de sélection de sous-systèmes adaptées à différents cas de figure (notamment à la présence de classes minoritaires). Ces stratégies résultent d'une analyse théorique de contrôle du conditionnement des sous-systèmes utilisés par l'algorithme RANSAC, ce qui permet d'améliorer l'efficacité de l'algorithme en le focalisant sur les sous-systèmes les plus stables numériquement.

Des simulations ont également permis d'évaluer empiriquement les performances de l'approche. En particulier, nous avons cherché les limites empiriques du domaine d'application de la méthode en terme de proportion du changement au sein des pixels basse résolution. Pour un niveau de contraste moyen, par exemple, les changements impactant plus de 25% d'un pixel basse

résolution sont bien détectés et ce tant que moins de 65% des pixels de l'image sont touchés.

Pour détecter des changements à partir de séquences d'images, il est nécessaire que les images soient superposables, *i.e.* préalablement recalées. En effet, plusieurs facteurs influent sur la géométrie du champ de vision acquis par un capteur satellite : la date de la prise de vue, l'angle de visée, le système de projection des données, la direction de balayage, la résolution spatiale, etc. Le recalage des images nécessite alors de replacer les images dans le même système de projection (UTM, Lambert NC, etc), de corriger les déformations dues à la trajectoire du satellite et de trouver la transformation qui relie chaque paire d'images. Même si ces corrections font l'objet de pré-traitements, des erreurs de recalage ont tendance à subsister et les méthodes de détection de changements sont généralement très sensibles à ces erreurs. Les travaux que nous avons présentés dans cette thèse supposent que les images sont parfaitement recalées entre elles. Cependant, nous avons réalisé quelques expériences dans l'objectif de tester la robustesse de la méthode de détection de changements en présence d'éventuels décalages. Une fois que les images sont dans le même système de projection et corrigées des déformations dues à la trajectoire du satellite, elles peuvent être rendues superposables par une combinaison de transformations, telles que rotations, translations ou homothéties. La définition du nombre de fausses alarmes dans le cadre de la détection de changements autorise la comparaison des valeurs de *NFA* obtenues à partir de données différentes. Dans un contexte expérimental, différentes transformations simples ont été appliquées sur des images **HR** avant de dégrader les images par moyennage par blocs. En considérant un ensemble fini de transformations, la recherche du *NFA* minimal sur cet ensemble de transformations permet de retrouver, à partir des images **BR**, le décalage simulé. Même si des expériences complémentaires sont nécessaires avant de pouvoir conclure, ces résultats annoncent une certaine robustesse de l'approche vis-à-vis des erreurs résiduelles de recalage. Après validation sur un plus grand nombre d'images ainsi que sur des données réelles, l'utilisation de la méthode de détection pourrait ainsi être étendue au cas de données mal recalées.

Pour finir, remarquons que l'applicabilité de nos méthodes à des situations réelles dépendra fortement de la précision de la correspondance entre les données haute et basse résolution, qu'elle soit donnée ou obtenue indirectement par l'introduction de paramètres de déformation de l'algorithme de classification. D'après l'étude réalisée, avec une bonne modélisation de la correspondance entre les images **HR** et **BR**, l'efficacité des méthodes proposées devrait être conservée. Les limites engendrées par la précision du recalage pourront être explorées pour préciser le domaine d'applicabilité de nos méthodes.



# Liste de publications

## Actes de conférence :

- *An unsupervised approach for subpixelic land-cover change detection.*  
A. Robin, L. Moisan, S. Le Hégarat-Masclé.  
Multitemp 2007, Fourth International Workshop on the Analysis of Multitemporal Remote Sensing Images, Leuven (Belgique), 18-20 Juillet 2007.
- *A multiscale land cover classification method using a Bayesian approach.*  
A. Robin, S. Le Hégarat-Masclé, L. Moisan.  
11<sup>th</sup> SPIE International Symposium on Remote Sensing, Bruges (Belgique), 19-22 Septembre 2005.
- *Approche a contrario pour la détection de changements à partir d'images basse résolution.*  
A. Robin, L. Moisan, S. Le Hégarat-Masclé.  
20<sup>e</sup> colloque GRETSI sur le traitement du signal et des images, Louvain-la-Neuve (Belgique),  
6-9 Septembre 2005.
- *Land cover classification from coarse resolution time series.*  
A. Robin, S. Le Hégarat-Masclé, L. Moisan  
31st International Symposium on Remote Sensing of Environment, Saint Petersburg (Russie),  
20-24 Juin 2005.
- *Traitement d'images satellites pour l'estimation de paramètres atmosphériques.*  
A. Robin, T. de Smit, J.-P. Berroir, I. Herlin, H. Yahia.  
Conférence ORASIS, Nancy 2003.

## Prépublications :

- *Automatic land cover change detection from coarse resolution images using an a contrario approach.*  
A. Robin, L. Moisan, S. Le Hégarat-Masclé.  
prépublication du MAP5 2005-3 (2005).



# Bibliographie





# Bibliographie

- [Agouris et al., 2000] Agouris, P., Gyftakis, S. et Stefanidis, A., Uncertainty in image-based change detection. In : Proc. of International Symposium on Spatial Accuracy, Amsterdam (Netherlands), p. 18. 2000.
- [Aiazzi et al., 1999] Aiazzi, B., Alparone, L., Barducci, A., Baronti, S. et Pippi, I., Multispectral fusion of multisensor image data by the generalized laplacian pyramid. In : IGARSS, Vol. 2, pp. 1183–1185. 1999.
- [Atkinson, 1997] Atkinson, P., Mapping sub-pixel boundaries from remotely sensed images. In : Z. Kemp (ed.), Innovations in GIS IV, Taylor and Francis, London, pp. 166–180. 1997.
- [Atkinson, 2001] Atkinson, P., Super-resolution target mapping from soft-classified remotely sensed imagery. In : Proceedings of the 6th International Conference on Geocomputation, Brisbane, University of Queensland. 2001.
- [Baret et Guyot, 1991] Baret, F. et Guyot, G., Potentials and limits of vegetation indices for lai and apar assessment. Remote Sensing of Environment 35, pp. 161–173. 1991.
- [Baret et al., 1989] Baret, F., Guyot, G. et Major, D., Tsavi : A vegetation index which minimizes soil brightness effects on lai and apar estimation. In : 12th Canadian symposium on remote sensing and IGARSS'90, Vol. 4, IEEE, pp. 1355–1359. 1989.
- [Bazi et Melgani, 2005] Bazi, Y., Bruzzone, L. et Melgani, F., An unsupervised approach based on the generalized gaussian model to automatic change detection in multitemporal SAR images. IEEE Transactions on Geoscience and Remote Sensing 43(4), pp. 874–887. 2005.
- [Benjamini et Hochberg, 1995] Benjamini, Y. et Hochberg, Y., Controlling the false discovery rate : A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 57(1), pp. 289–300. 1995.
- [Biernacki, 1999] Biernacki, C., Précision sur les données et coude de la vraisemblance pour trouver le nombre de classes dans un mélange. Revue de Statistique Appliquée 47(1), pp. 47–62. 1999.
- [Binaghi et al., 1999] Binaghi, E., Brivio, P., Ghessi, P. et Rampini, A., A fuzzy set based accuracy assessment of soft classification. Pattern Recognition Letters 20, pp. 935–948. 1999.
- [Bonferroni, 1936] Bonferroni, C., Teoria statistica delle classi et calcolo delle probabilita. Pubblicazioni del Istituto Superiore de Scienze Economiche e Commerciali di Firenze 8, pp. 3–62. 1936.
- [Borak et al., 2000] Borak, J., Lambin, E. et Strahler, A., The use of temporal metrics for land cover change detection at coarse spatial scales. International Journal of Remote Sensing 21(6-7), pp. 1415–1432. 2000.
- [Bosc et al., 2003] Bosc, M., Heitz, F., Armspach, J., Namer, I., Gounot, D. et Rumbach, L., Automatic change detection in multimodal serial mri : application to multiple sclerosis lesion evolution. Neuroimage 20, pp. 643–656. 2003.

- [Boucher et Kyriakidis, 2006] Boucher, A. et Kyriakidis, P., Super-resolution land cover mapping with indicator geostatistics. *Remote Sensing of Environment* 104, pp. 264–282. 2006.
- [Brown et al., 2000] Brown, M., Lewis, H. et Gunn, S., Linear spectral mixture models and support vector machines remote sensing. *IEEE Transactions on Geoscience and Remote Sensing* 38, pp. 2346–2360. 2000.
- [Bruzzone et Prieto, 2000] Bruzzone, L. et Prieto, D., An adaptative parcel-based technique for unsupervised change detection. *International Journal of Remote Sensing* 21(4), pp. 817–822. 2000.
- [Bruzzone et Prieto, 2002] Bruzzone, L. et Prieto, D., An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 11(4), pp. 452–466. 2002.
- [Cao et al., 2005] Cao, F., Veit, T. et Bouthemy, P., Image comparison and motion detection by a contrario methods. In : L. Harris et M. Jenkin (eds), *Computational Vision in Neural and Machine Systems*, Cambridge University Press. 2005.
- [Cardot et al., 2003] Cardot, H., Faivre, R. et Goulard, M., Functional approaches for predicting land use with the temporal evolution of coarse resolution remote sensing data. *Journal of Applied Statistics* 30, pp. 1185–1199. 2003.
- [Carincotte et al., 2004] Carincotte, C., Derrode, S. et Bourennane, S., Unsupervised image segmentation based on high-order hidden markov chains. In : *Proc. of IEEE International Conference on Acoustics, Speech, Signal Processing*, Montreal (Canada). 2004.
- [Carper et al., 1990] Carper, W., Lillesand, T. et Kiefer, R., The use of intensity-hue-saturation transformations for merging spot panchromatic and multispectral image data. *Photogrammetric Engineering Remote Sensing* 56(4), pp. 459–467. 1990.
- [Catoni, 1998] Catoni, O., Solving scheduling problems by simulated annealing. *SIAM J. Control Optim.* 36(5), pp. 1539–1575 (electronic). 1998.
- [CCT, 2001] CCT, Notions fondamentales de télédétection. Rapport de recherche, Centre Canadien de Télédétection, [http://cct.rncan.gc.ca/resource/tutor/fundam/index\\_f.php](http://cct.rncan.gc.ca/resource/tutor/fundam/index_f.php). 2001.
- [Celeux et Diebolt, 1986] Celeux, G. et Diebolt, J., L'algorithme sem : un algorithme d'apprentissage probabiliste pour la reconnaissance de mélange de densités. *Revue de Statistique Appliquée* 34(2), pp. 35–52. 1986.
- [Chandrasekhar, 1950] Chandrasekhar, S., *Radiative Transfer*. New-York. 1950.
- [Chavez et Anderson, 1991] Chavez, P.S. and Sides, S. et Anderson, J., Comparison of three different methods to merge multiresolution and multispectral data : LandSat TM and Spot panchromatic. *Photogrammetric Engineering Remote Sensing* 57(3), pp. 295–303. 1991.
- [Ciarlet, 1982] Ciarlet, P. G., *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree], Masson, Paris. 1982.
- [Collins et Woodcock, 1996] Collins, J. et Woodcock, C., An assessment of several linear change detection techniques for mapping forest mortality using multitemporal landsat tm data. *Remote Sensing of Environment* 56, pp. 66–77. 1996.
- [Cooper et al., 1982] Cooper, K., Smith, J. et Pitts, D., Reflectance of a vegetation canopy using the adding method. *Applied Optics* 21(22), pp. 4112–4118. 1982.

- 
- [Coppin et Bauer, 1994] Coppin, P. et Bauer, M., Processing of multitemporal landsat tm imagery to optimise extraction of forest cover change features. *IEEE Transactions on Geoscience and Remote Sensing* 32, pp. 918–927. 1994.
- [Coppin et al., 2004] Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B. et Lambin, E., Digital change detection methods in ecosystem monitoring : a review. *International Journal of Remote Sensing* 25(9), pp. 1565–1596. 2004.
- [Cross et al., 1991] Cross, A., Settle, J., Drake, N. et Paivinen, R., Subpixel measurement of tropical forest cover using avhrr data. *International Journal of Remote Sensing* 12, pp. 1119–1129. 1991.
- [Cutler et Windham, 1993] Cutler, A. et Windham, M., Information-based validity functionals for mixture analysis. *Proceedings of the first US-Japan Conference on the Frontiers of Statistical Modeling*, Amsterdam, pp. 149–170. 1993.
- [Dai et Khorram, 1998] Dai, X. et Khorram, S., The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing* 36(5), pp. 1566–1577. 1998.
- [Deer et Eklund, 2002] Deer, P. et Eklund, P., *Fuzzy logic for change detection in classified images*. Springer Verlag, Berlin. 2002.
- [Dempster et al., 1977] Dempster, A., Laird, N. et Rubin, D., Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B.* 39(1), pp. 1–38. 1977.
- [Desolneux et al., 2000] Desolneux, A., Moisan, L. et Morel, J., Meaningful alignments. *Int. J. Comp. Vision* 40(1), pp. 7–23. 2000.
- [Desolneux et al., 2001] Desolneux, A., Moisan, L. et Morel, J., Edge detection by helmholtz principle. *J. Math. Imaging and Vision* 14(3), pp. 271–284. 2001.
- [Desolneux et al., 2003] Desolneux, A., Moisan, L. et Morel, J., A grouping principle and four applications. *IEEE Transactions on Pattern and Machine Intelligence* 25(4), pp. 508–513. 2003.
- [Desolneux et al., 2006] Desolneux, A., Moisan, L. et Morel, J., From gestalt theory to image analysis - a probabilistic approach. preprint MAP5 2006-9, 2006.
- [Dibos et al., 2005] Dibos, F., Pelletier, S. et Koepfler, G., Real-time segmentation of moving objects in a video sequence by a contrario motion detection. In : *Proc. of IEEE International Conference on Image Processing*, Genova, Italy. 2005.
- [Duggin, 1985] Duggin, M., Factors limiting the detection and quantification of terrestrial features using remotely sensed radiance. *International Journal of Remote Sensing* 6, pp. 3–27. 1985.
- [Edwards, 1990] Edwards, G., Image segmentation, cartographic information and knowledge-based reasoning : getting the mixture right. In : *Proc. of the International Geosciences and Remote Sensing Symposium*, pp. 1641–1644. 1990.
- [Eklund et al., 2000] Eklund, P., You, J. et Deer, P., Mining remote sensing image data : an integration of fuzzy set theory and image understanding techniques for environmental change detection. In : *Proceedings of SPIE - Data Mining and Knowledge Discovery : Theory, Tools, and Technology*, Orlando, Florida (USA), pp. 265–272. 2000.
- [Faivre et al., 2000] Faivre, R., Bastié, C. et Husson, A., Integration of vegetation and hrvir data into yield estimation approach. In : *Proceedings of Vegetation 2000, 2 years of Operation to Prepare the Future.*, Joint Research Center, Ispra, Varese (Italie), pp. 235–240. 2000.
-

- [Faivre et Fischer, 1997] Faivre, R. et Fischer, A., Predicting crop reflectances using satellite data observing mixed pixels. *Journal of Agricultural, Biological and Environmental Statistics* 2, pp. 87–107. 1997.
- [Filiberti et al., 1994] Filiberti, D., Marsh, S. et Schowengerdt, R., Synthesis of imagery with high spatial and spectral resolution from multiple image sources. *Optical Engineering* 33(8), pp. 2520–2528. 1994.
- [Fischler et Bolles, 1981] Fischler, M. et Bolles, R., Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24, pp. 381–385. 1981.
- [Fisher, 1994] Fisher, A., A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote Sensing and environment*. 1994.
- [Foody, 1996] Foody, G., Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data. *International Journal of Remote Sensing* 17, pp. 1317–1340. 1996.
- [Foody, 1998] Foody, G., Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution. *International Journal of Remote Sensing* 19(13), pp. 2593–2599. 1998.
- [Foody, 2000] Foody, G., Mapping land cover from remotely sensed data with a softened feedforward neural network. *Journal of Intelligent and Robotic System* 29, pp. 433–449. 2000.
- [Fung, 1990] Fung, T., An assessment of tm imagery for land-cover change detection. *IEEE Transactions on Geoscience and Remote Sensing* 28(12), pp. 681–684. 1990.
- [Fung et Le Drew, 1987] Fung, T. et Le Drew, E., Application of principal component analysis to change detection. *Photogrammetric Engineering & Remote Sensing* 53, pp. 1649–1658. 1987.
- [Geman et Geman, 1984] Geman, D. et Geman, S., Stochastic relaxation, gibbs distribution, and the bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 6, pp. 721–741. 1984.
- [Gong, 1993] Gong, P., Change detection using principal components analysis and fuzzy set theory. *Canadian Journal of Remote Sensing*. 1993.
- [Grosjean et Moisan, 2006] Grosjean, B. et Moisan, L., A-contrario detectability of spots in textured backgrounds. MAP5 pre-publication, 2006.
- [Hanaizumi et Fujimura, 1992] Hanaizumi, H. et Fujimura, S., Change detection from remotely sensed multi-temporal images using multiple regression. In : *Proceedings from the 1992 International Geoscience and Remote Sensing Symposium, IGARSS, Houston*, pp. 564–566. 1992.
- [Hochberg, ] Hochberg, Y., A sharper bonferroni procedure for multiple tests of significance. *Biometrika*.
- [Hochberg et Tamhane, 1987] Hochberg, Y. et Tamhane, A., *Multiple comparison procedures*. John Wiley & Sons, New-York. 1987.
- [Hommel, 1988] Hommel, G., A stagewise rejective multiple test procedure based on a modified bonferroni test. *Biometrika* 75, pp. 383–386. 1988.
- [Horwitz et al., 1971] Horwitz, H., Nalepka, R., Hyde, P. et Morgenstern, J., Estimating the proportions of objects within a single resolution element of a multispectral scanner. In : *Proceedings of the 7th International Symposium on Remote Sensing of Environment.*, Ann Arbor, Michigan, pp. 1307–1320. 1971.
- [Huete, 1988] Huete, A., A soil-adjusted vegetation index (savi). *Remote Sensing of Environment* 25, pp. 295–309. 1988.
- [Inglada, 2001] Inglada, J., *Etat de l’art en détection de changements sur les images de télédétection*. Note technique, CNES. 2001.

- 
- [Jacquemoud et al., 1995] Jacquemoud, S., Baret, F., Andrieu, B., Danson, F. et Jaggard, K., Extraction of vegetation biophysical parameters by inversion of the prospect+sail models on sugar beet canopy reflectance data. application to tm and aviris sensors. *Remote Sensing of Environment* 52(3), pp. 163–172. 1995.
- [Jensen et Cowen, 1997] Jensen, J. et Cowen, D., Principles of change detection using digital remote sensor data. In : M. Estes (ed.), *Integration of Geographic Information System and Remote Sensing*, Cambridge Press, London (England), pp. 37–54. 1997.
- [Justice et Townshend, 2002] Justice, C. et Townshend, J., Special issue on the moderate resolution imaging spectroradiometer (modis) : a new generation of land surface monitoring. *Remote Sensing of Environment*. 2002.
- [Kasetkasem et al., 2005] Kasetkasem, T., Arora, M. et Varshney, P., Super-resolution land cover mapping using a markov random field based approach. *Remote Sensing of Environment* 96, pp. 302–314. 2005.
- [Kasetkasem et Varshney, 2002] Kasetkasem, T. et Varshney, P., An image change detection algorithm based on markov random field models. *IEEE Transactions on Geoscience and Remote Sensing* 40(8), pp. 1815–1823. 2002.
- [Kay, 1993] Kay, S., *Fundamentals of Statistical Signal Processing : Detection Theory*. Prentice-Hall. 1993.
- [Kerdiles et Grondona, 1995] Kerdiles, H. et Grondona, M., Noaa-avhrr ndvi decomposition and subpixel classification using linear mixing in the argentinean pampa. *International Journal of Remote Sensing* 16(7), pp. 1303–1325. 1995.
- [Knipling, 1970] Knipling, E., Physical and physiological bases for the reflectance of visible and near infrared radiation from vegetation. *Remote Sensing of Environment* 1, pp. 155–159. 1970.
- [Koepfler et al., 1994] Koepfler, G., Lopez, C. et Morel, J., A multiscale algorithm for image segmentation by variational method. *SIAM Journal on Numerical Analysis* 31(1), pp. 282–299. 1994.
- [Konrad, 2000] Konrad, J., *Handbook of Image and Video Processing*. Academic Press, chapter Motion detection and estimation, pp. 207–225. 2000.
- [Lambin et Stralher, 1994] Lambin, E. et Stralher, A., Change vector analysis in multitemporal space : a tool to detect and categorize land-cover change processes using high temporal-resolution satellite data. *Remote Sensing of Environment* 48, pp. 231–244. 1994.
- [Laporterie-Déjean et al., 2003] Laporterie-Déjean, F., Lopez-Ornelas, E. et Flouzat, G., Pre-segmentation of high-resolution images thanks to the morphological pyramid. In : *Proc. of Int. Geoscience and Remote Sensing Symposium, IGARSS'03*, Vol. 3, pp. 2042–2044. 2003.
- [Le Hégarat-Masclé et Seltz, 2004] Le Hégarat-Masclé, S. et Seltz, R., Automatic change detection by evidential fusion of change indices. *Rem. Sens. Environment* 91, pp. 390–404. 2004.
- [Le Hégarat-Masclé et al., 2005] Le Hégarat-Masclé, S., Ottlé, C. et Guérin, C., Land cover change detection at coarse spatial scales based on iterative estimation and previous state information. *Remote Sensing of Environment* 95, pp. 464–479. 2005.
- [Lee, 2003] Lee, Y., A study on accuracy of georeferencing of multisensor remote sensing images. In : *Proceedings of International Symposium on Spatial Data Quality*, Hong-Kong, China. 2003.
- [Liu, 2001] Liu, J., *Monte-Carlo strategies in scientific computing*. Springer-Verlag, New-York. 2001.
- [Lowe, 1985] Lowe, D., *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers. 1985.
-

- [Lu et al., 2004] Lu, D., Mausel, P., Brondizio, E. et Moran, E., Change detection techniques. *International Journal of Remote Sensing* 25(12), pp. 2365–2407. 2004.
- [Lyon et al., 1998] Lyon, J., Yuan, D., Lunetta, R. et Elvidge, C., A change detection experiment using vegetation indices. *Photogrammetric Engineering and Remote Sensing* (64), pp. 143–150. 1998.
- [Makido et Shortridge, 2005] Makido, Y. et Shortridge, A., Land cover mapping at subpixel scales : Unraveling the mixel pixel. In : *Proceedings of the 8th International Conference on Geocomputation*, Ann Arbor, University of Michigan. 2005.
- [Malila, 1980] Malila, W., Change vector analysis : an approach for detecting forest changes with landsat. In : *Proc. of the Annual Symposium on Machine Processing of Remotely Sensed Data*, pp. 326–335. 1980.
- [Mas, 1999] Mas, J., Monitoring land-cover changes : a comparison of change detection techniques. *International Journal of Remote Sensing* 20(1), pp. 139–152. 1999.
- [Mertens et al., 2003] Mertens, K., Verbeke, L., Ducheyne, E. et De Wulf, R., Using genetic algorithms in sub-pixel mapping. *International Journal of Remote Sensing* 24(21), pp. 4241–4247. 2003.
- [Mertens et al., 2004] Mertens, K., Verbeke, L., Westra, T. et De Wulf, R., Sub-pixel mapping and sub-pixel sharpening using neural network predicted wavelet coefficients. *Remote Sensing of Environment* 91(2), pp. 225–236. 2004.
- [Moisan et Stival, 2004] Moisan, L. et Stival, B., A probabilistic criterion to detect rigid point matches between two images and estimate the fundamental matrix. *Int. J. Comp. Vision* 57(3), pp. 201–218. 2004.
- [Moser et Serpico, 2006] Moser, G. et Serpico, S., Generalized minimum-error thresholding for unsupervised change detection from sar amplitude imagery. *IEEE Transactions on Geoscience and Remote Sensing* 44(10), pp. 2972–2982. 2006.
- [Mumford et Shah, 1989] Mumford, D. et Shah, J., Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.* 42, pp. 577–685. 1989.
- [Neyman et Pearson, 1967] Neyman, J. et Pearson, E., The testing of statistical hypotheses in relation to probabilities a priori. *Joint Statistical Papers*, Cambridge university press (originally printed in 1933) pp. 186–202. 1967.
- [Nielsen, 2005] Nielsen, A., An iterative extension to the MAD transformation for change detection in multi- and hyperspectral remote sensing data. In : *4th EARSeL Workshop on Imaging Spectroscopy*, Technical University of Denmark. 2005.
- [Nielsen et al., 1998] Nielsen, A., Conradsen, K. et Simpson, J., Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data : New approaches to change detection studies, remote sensing of environment. *Remote Sensing of Environment* 64, pp. 1–19. 1998.
- [O'Sullivan, 1993] O'Sullivan, F., Mixture estimation from multichannel image data. *Journal of the American Statistical Association* 88(421), pp. 209–220. 1993.
- [Pearson et Miller, 1972] Pearson, R. et Miller, L., Remote mapping of standing crop biomass for estimation of the productivity of the short grass prairie. In : A. A. Ed. (ed.), *8th International Symposium on Remote Sensing of the Environment*, Pawnee National Grasslands, Colorado, pp. 1357–1381. 1972.
- [Pieczynski, 1992] Pieczynski, W., Statistical image segmentation. *Machine Graphics & Vision* 1(1/2), pp. 261–268. 1992.

- 
- [Pohl et Van Genderen, 1998] Pohl, C. et Van Genderen, J., Multisensor image fusion in remote sensing : concepts, methods and application. *Int. Journal of Remote Sensing* 19(5), pp. 823–854. 1998.
- [Press et al., 1988] Press, W., Teukolsky, S., Vetterling, W. et Flannery, B., *Numerical Recipes* in C. Cambridge University Press. 1988.
- [Qi et al., 1994] Qi, J., Chehbouni, A., Huete, A. et Kerr, Y., A modified soil-adjusted vegetation index. *Remote Sensing of Environment* 48, pp. 119–126. 1994.
- [Radke et al., 2005] Radke, R., Andra, S., Al-Kohafi, O. et Roysam, B., Image change detection algorithms : A systematic survey. *IEEE Transactions on Image Processing* 14(3), pp. 294–307. 2005.
- [Ranchin et Wald, 2000] Ranchin, T. et Wald, L., Fusion of high spatial images : the ARSIS concept and its implementation. *Photogrammetric Engineering and Remote Sensing* 66(1), pp. 49–61. 2000.
- [Rey et al., 2002] Rey, D., Subsol, G., Delingette, H. et Ayache, N., Automatic detection and segmentation of evolving processes in 3d medical images : Application to multiple sclerosis. *Medical Image Analysis* 6(2), pp. 163–179. 2002.
- [Richardson et Wiegand, 1977] Richardson, A. et Wiegand, C., Distinguishing vegetation from soil background information. *Photogrammetric Engineering Remote Sensing* 43, pp. 1541–1552. 1977.
- [Rignot et van Zyl, 1993] Rignot, E. et van Zyl, J., Change detection techniques for ERS-1 SAR data. *IEEE Transactions on Geoscience and Remote Sensing* 31(4), pp. 896–906. 1993.
- [Rosin, 2002] Rosin, P., Thresholding for change detection. *Computer Vision and Image Understanding* 86(2), pp. 79–95. 2002.
- [Rouse et al., 1973] Rouse, J. W., Haas, R., Schell, J. et Deering, D., Monitoring vegetation systems in the great plains with erts. In : *Third ERTS Symposium*, NASA SP-351 I, pp. 309–317. 1973.
- [Rubinstein, 1981] Rubinstein, R., *Simulation and the Monte-Carlo method*. Wiley, New-York. 1981.
- [Salvi et al., 2001] Salvi, J., Armangué, X. et Pagès, J., A survey addressing the fundamental matrix estimation problem. In : *IEEE Int. Conf. on Image Processing*, Thessaloniki (Greece). 2001.
- [Saporta, 1990] Saporta, G., *Probabilités, Analyse des Données et Statistique*. Editions TECHNIP. 1990.
- [Settle et Campbell, 1998] Settle, J. et Campbell, N., On the errors of the two estimators of subpixel fractinal cover when mixing is linear. *IEEE Transactions on Geoscience and Remote Sensing* 36, pp. 163–170. 1998.
- [Settle et Drake, 1993] Settle, J. et Drake, N., Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing* 14(6), pp. 1159–1177. 1993.
- [Singh, 1989] Singh, A., Digital image change detection techniques using remotely-sensed data. *International Journal of Remote Sensing* 10(6), pp. 989–1003. 1989.
- [Small, 2001] Small, C., Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing* 22(7), pp. 1305–1334. 2001.
- [Smits et Annoni, 2000] Smits, P. et Annoni, A., Toward specification-driven change-detection. *IEEE Transactions on Geoscience and Remote Sensing* 38(3), pp. 1484–1488. 2000.
- [Sveinsson et Benediktsson, 2000] Sveinsson, J. et Benediktsson, J., Data fusion and feature extraction using tree structured filter banks. In : *IGARSS*, Vol. 6, pp. 2617–2619. 2000.
- [Tatem et al., 2001] Tatem, A., Lewis, H., Atkinson, P. et Nixon, M., Super-resolution target identification from remotely sensed images using a hopfield neural network. *IEEE Transactions on Geoscience and Remote Sensing* 39(4), pp. 781–796. 2001.
-



- [Tatem et al., 2002] Tatem, A., Lewis, H., Atkinson, P. et Nixon, M., Super-resolution land cover pattern prediction using a hopfield neural network. *Remote Sensing of Environment* 79(1), pp. 1–14. 2002.
- [Thévenaz et Unser, 2000] Thévenaz, P. et Unser, M., Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing* 9(12), pp. 2083–2099. 2000.
- [Torr et Murray, 1997] Torr, P. et Murray, D., The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision* 24(3), pp. 271–300. 1997.
- [Tseng et al., 2001] Tseng, D., Chen, Y. et Liu, M., Wavelet-based multispectral image fusion. In : *IGARSS*, Vol. 4, pp. 1956–1958. 2001.
- [Uenishi et al., 2005] Uenishi, T., Oki, K., Omasa, K. et Tamura, M., A land cover distribution composite image from coarse spatial resolution images using an unmixing method. *International journal of Remote Sensing* 26(5), pp. 871–886. 2005.
- [Van Der Meer, 1999] Van Der Meer, F., Iterative spectral unmixing. *International Journal of Remote Sensing* 20(17), pp. 3431–3436. 1999.
- [Verhoef, 1984] Verhoef, W., Light scattering by leaf layers with application to canopy reflectance modeling : the sail model. *Remote Sensing of Environment* 16, pp. 125–141. 1984.
- [Verhoeve et De Wulf, 2002] Verhoeve, J. et De Wulf, R., Land cover mapping at sub-pixel scales using linear optimization techniques. *Remote Sensing of Environment* 79, pp. 96–104. 2002.
- [Vignolles, 1996] Vignolles, C., Modélisation de la production agricole à moyenne échelle - Apport du couplage des mesures par satellite à haute résolution spatiale et à haute résolution temporelle. Thèse, Université Paul Sabatier. 1996.
- [Weydahl, 1991] Weydahl, D., Change detection in SAR images. In : *Proc. of IEEE International Conference on Geoscience and Remote Sensing*, pp. 1421–1424. 1991.
- [Wiemker et al., 1997] Wiemker, R., Speck, A., Kulbach, D., Spitzer, H. et Bienlein, J., Unsupervised robust change detection on multispectral imagery using spectral and spatial features. In : *Proc. of International Airborne Remote Sensing Conference and Exhibition*, Copenhagen, Denmark. 1997.
- [Wren et al., 1997] Wren, C., Azarbayejani, A., Darrell, T. et Pentland, A., Pfunder : Real-time tracking of the human body. *IEEE Transaction Pattern Analysis and Machine Intelligence* 19(7), pp. 780–785. 1997.
- [Yasuoka, 1988] Yasuoka, Y., Detection of land-cover changes from remotely sensed images using spectral signature similarity. In : *Proceedings of the 9th Asian Conference on Remote Sensing*, pp. G31–G36. 1988.
- [Ycart, 2002] Ycart, B., Modèles et algorithmes markoviens. *Mathématiques & Applications (Berlin) [Mathematics & Applications]*, Vol. 39, Springer-Verlag, Berlin. 2002.
- [Zhang et al., 1994] Zhang, Z., Deriche, R., Faugeras, . et Luong, Q.-T., A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *AI Journal* (78), pp. 87–119. 1994.
- [Zhukov et al., 1999] Zhukov, B., Oertel, D., Lanzl, F. et Reinhäckel, G., Unmixing-based multisensor multiresolution image fusion. *IEEE Trans. Geo. Rem. Sens.* 37(3), pp. 1212–1226. 1999.

# Table des figures

1.1. Représentation schématique de l'orbite d'un satellite géostationnaire (à gauche) et de la fauchée d'un satellite en orbite quasi-polaire (à droite) pour l'observation de la Terre( <i>cf.</i> [CCT, 2001]). . . . .	19
1.2. L'image (a) représente un extrait d'une image de réflectance <b>SPOT/HRV</b> de la plaine du Danube acquise en Juin 2001. Les images (b) et (c) représentent des images du même site simulées par moyennage par blocs respectivement de taille $16 \times 16$ (moyenne résolution, telle que celle de <b>MeRIS</b> ) et de taille $50 \times 50$ (basse résolution, telle que celle de <b>SPOT/VGT</b> ). . . . .	20
1.3. Interactions entre la couche de végétation et la surface terrestre. Les rayons lumineux notés $R_v$ , $R_s$ et $T_v$ désignent, respectivement, les rayons réfléchis par la végétation, par le sol et transmis par la végétation. . . . .	26
1.4. Les réponses spectrales de la végétation ( <i>cf.</i> [CCT, 2001]) : mesure de réflectance en fonction de la longueur d'ondes (en nanomètres). . . . .	27
1.5. Profils temporels de fractions de couvert : le temps est représenté en abscisse et la fraction de couvert en ordonnée. Chaque graphe représente l'évolution temporelle de la fraction de couvert observée en un pixel de l'image. . . . .	31
5.1. Majorant de la probabilité de préférer une labellisation 1-optimale à la labellisation correcte lorsque la région mal labellisée représente une proportion supérieure à $n/N$ d'un pixel basse résolution. La valeur du majorant est tracée en fonction du niveau <i>contraste</i> de l'image. L'axe des abscisses représente le niveau de contraste $c$ et l'axe des ordonnées le majorant de la probabilité. Chaque courbe correspond à une proportion $n/N$ fixée, allant de 0.4% à 74.6%. . . . .	83
6.1. Entrées et sortie de l'algorithme : à partir d'une segmentation (300 régions) et d'une image de même résolution, l'algorithme permet d'obtenir une classification pour un nombre de labels fixé (ici, $L = 8$ ). . . . .	87
6.2. Exemples de labellisations simulées avec 5 labels. . . . .	90
6.3. Images <b>BR</b> simulées à partir des labellisations ci-dessus, avec facteurs de résolution, de gauche à droite, $N = 15, 30, 50$ . . . . .	90

6.4. Labellisation pour 10 labels de la segmentation présentée en (a) à partir de l'image <b>BR</b> (b) (rapport de résolution $N = 15$ ), puis à partir d'une séquence basse résolution contenant 2 images ((b) et une autre image similaire), et enfin à partir de 3 images. Les erreurs de labellisation diminuent lorsqu'on augmente le nombre d'images exploitées. . . . .	92
6.5. Labellisation pour 10 labels de la segmentation présentée en (a) à partir de la séquence <b>BR</b> composée des images (b) et (c) (rapport de résolution $N = 15$ ), puis à partir des images (b), (c) et (d). Les erreurs de labellisation diminuent avec lorsque des informations redondantes sont utilisées. . . . .	93
6.6. Labellisation obtenue en mode non-supervisé, avec une séquence de 5 images <b>BR</b> , simulées à partir des profils temporels (a) et avec un facteur de résolution de 15 par rapport à la résolution de la segmentation (g). La carte des erreurs (i) montre que 5 régions ont été mal labellisées, elles représentent 0.28% des pixels de l'image. . . . .	94
6.7. (a) Histogramme de la différence d'énergies normalisée $(\mathcal{E}_k(v) - \mathcal{E}_0(v))/\mathcal{E}_0(v)$ calculée pour toutes les régions de la segmentation labellisées successivement avec tous les labels incorrects. (b) Histogramme de la différence d'énergies normalisée $(\mathcal{E}_k(v) - \mathcal{E}_0(v))/\mathcal{E}_0(v)$ calculée pour toutes les erreurs obtenues avec l'algorithme supervisé. La comparaison de ces histogrammes montre, notamment, que les erreurs de labellisation dues à un problème de convergence sont associées à des différences d'énergies très faibles par rapport à celles qu'on obtiendrait avec d'autres labels inexactes). . . . .	96
6.8. Histogramme du taux d'occupation des régions mal labellisées au sein des pixels basse résolution pour différents facteurs de résolution. L'axe des abscisses représente le taux d'occupation d'une région mal labellisée au sein d'un pixel basse résolution et l'axe des ordonnées représente le pourcentage d'occurrence. . . . .	98
6.9. Statistiques du taux d'occupation d'une région mal labellisée dans un pixel basse résolution en fonction du facteur de résolution considéré ( $N$ ). . . . .	98
6.10. Chaque croix représente le pourcentage des pixels mal labellisés obtenus pour un jeu de données et un rapport de résolution fixé avec la version supervisée (a) et avec la version non-supervisée (b). Le rapport de résolution $N$ est représenté en abscisses et le pourcentage en ordonnées. La ligne verte superposée correspond à la médiane de ces valeurs de pourcentage. . . . .	99
6.11. Pourcentage des pixels mal labellisés obtenus avec l'algorithme supervisé, à partir d'images réelles monotemporelles. . . . .	100
8.1. A gauche, la labellisation <b>HR</b> contenant 9 labels (entrée de l'algorithme). A droite, image <b>HR</b> simulée à partir de la labellisation en introduisant 2 zones de changements de type différents. . . . .	139

8.2. A gauche, en haut : image <b>BR</b> simulée par moyennage de l'image figure 8.1(b) par blocs de taille $16 \times 16$ ; en bas, les pixels de changements détectés (27) sont représentés en noir, et la trace des contours des changements introduits à haute résolution est superposée en blanc (la taille des pixels <b>BR</b> a été agrandie pour plus de lisibilité). A droite, les croix rouges représentent l'erreur quadratique minimale obtenue en fonction de la taille du sous-domaine $D$ étudié (échelle $\log_{10}$ ); en vert, les pointillés représentent la valeur minimale du $NFA$ obtenue en fonction de la taille du sous-domaine considéré. . . . .	140
8.3. Erreur quadratique (en échelle log) obtenue pour l'estimation des moyennes en fonction du nombre d'itérations réalisées ( $10^5$ ). Résultat obtenu avec tirage aléatoire des pixels. . . . .	141
8.4. Histogrammes du conditionnement pour la norme 2 obtenu pour 100 000 matrices $M$ de taille $6 \times 6$ tirées aléatoirement (échelle log). Tirage uniforme de matrices quelconques dans $[0, 1]$ figure (a), de matrices de proportions simulées figure (b) (terme général de la forme $\frac{a}{15}$ , $a \in [0, 15]$ ), et de matrices de proportions extraites aléatoirement à partir d'un jeu de données figure (c). . . . .	143
8.5. Valeur du $NFA$ obtenu en fonction du conditionnement de la matrice $A$ utilisée pour l'estimation des moyennes des classes, en échelle log. Résultats obtenus pour 100 000 itérations où les matrices sont sélectionnées aléatoirement (figure (a)). Figure (b) : valeurs médianes et des 10ème et 90ème percentiles obtenus sur la restriction de l'échantillon aux matrices numériquement inversibles, <i>i.e.</i> telles que $\text{cond}_2(A) \leq 1.e^6$ . Figure (c) : l'histogramme de fréquence 2D du couple $(\text{cond}_2(A), NFA)$ , en échelle log. . . . .	144
8.6. Histogrammes des valeurs de $NFA$ obtenues pour un intervalle de valeurs du conditionnement de $A$ donné. Chaque intervalle est défini de manière à compter le même nombre d'occurrences (1904). Plus le conditionnement de $A$ est grand (en échelle log), plus les valeurs du $\log(NFA)$ sont grandes. . . . .	145
8.7. Histogramme du conditionnement (échelle log) pour la norme 2 obtenu pour 100 000 matrices $A$ de taille $6 \times 6$ correspondant réellement à des proportions et sélectionnées, figure (a), aléatoirement (uniformément), figure (b) avec la stratégie 1 (seuil=0.2) et, figure (c), avec la stratégie 2 pour $n = 20$ . . . . .	148
8.8. Histogramme du conditionnement (échelle log) pour la norme 2 obtenu pour 100 000 matrices $A$ de taille $6 \times 6$ correspondant réellement à des proportions et sélectionnées selon la stratégie 2, avec $n = 5$ figure (a), avec $n = 10$ figure (b) et avec $n = 15$ figure (c). . . . .	151
8.9. Histogramme du conditionnement (échelle log) pour la norme 2 obtenu pour 100 000 matrices $A$ de taille $6 \times 6$ correspondant réellement à des proportions et sélectionnées selon la stratégie 3, avec $n = 5$ figure (a), avec $n = 10$ figure (b) et avec $n = 15$ figure (c). . . . .	152

9.1.	A gauche, chaque courbe représente la fonction $g$ en fonction de $p$ pour différentes valeurs de $\gamma$ . A droite, la valeur $p^*$ obtenue numériquement est tracée en fonction du paramètre $\gamma$ . . . . .	160
9.2.	Evolution de l'écart-type en fonction du facteur de résolution. Ecart-type calculé sur des images de fraction de couvert de la série ADAM (échelle SPOT), et sur leurs dégradations spatiales pour des facteurs allant de 1 à 80. Les courbes sont relativement superposables. Image acquise le 15/10/2000 en marron (jaune), le 31/10/2000 en rouge, le 22/02/2001 en turquoise, le 18/04/2001 en vert, le 06/06/2001 en bleu et le 05/07/2001 en mauve. . . . .	163
10.1.	Images <b>HR</b> simulées par tirage gaussien à partir de la labellisation (a), avec un écart-type $\sigma = 0.05$ (b) et $\sigma = 0.1$ (c). L'image <b>BR</b> (d) est obtenue par moyennage de l'image <b>HR</b> (b) par blocs de taille $16 \times 16$ . . . . .	168
10.2.	Pourcentage des pixels de l'image détectés à tort comme cohérents avec la labellisation (faux positifs, en bleu), de ceux détectés à tort comme changements (faux négatifs, en vert), et le total des erreurs obtenues en rouge. Valeur médiane des résultats obtenus sur 500 tests pour chaque pourcentage de bruit impulsionnel donné, avec des images simulées avec un écart-type $\sigma = 0$ (a), $\sigma = 0.05$ (b) et $\sigma = 0.1$ (c). Même avec un faible niveau de contraste, la méthode permet de détecter les pixels de changement avec moins de 10% d'erreurs en présence de 70% de bruit impulsionnel. . . . .	171
10.3.	Erreurs de détection obtenues en moyenne en fonction du pourcentage de changements introduit. Le pourcentage moyen de pixels détectés comme changements à tort par rapport au nombre pixels sans changements est représenté figure (a) et, figure (b), le pourcentage moyen de pixels détectés comme cohérents avec la classification à tort par rapport au nombre pixels de changements. Le taux de faux négatif est très robuste au changement de contraste, même en présence de 75% de pixels de changement et le taux de faux positifs est très robuste au pourcentage de pixels de changement dans l'image (tant qu'il est inférieur à 75%). . . . .	173
10.4.	Evolution des performances en fonction du nombre de pixels de changements présents dans l'image. A gauche, les points rouge sur la diagonale représentent les pixels de non-changements présents en réalité dans l'image testée, le nombre de pixels détectés comme changements à tort (faux négatifs) est représenté en bleu, et le nombre de pixels détectés comme valides à raison (vrais positifs) en vert. A droite, les points rouge sur la diagonale représentent les pixels de changements, le nombre de pixels valides à tort est représenté en croix bleues et, en vert, le nombre de pixels détectés comme changements. . . . .	175
10.5.	Images simulées avec un écart-type nul. . . . .	176
10.6.	Images simulées avec un écart-type $\sigma = 0.05$ . . . . .	176
10.7.	Images simulées avec un écart-type $\sigma = 0.1$ . . . . .	176

10.8. Les images (a) et (c) font apparaître en rouge, respectivement, les 20% et 40% de bruit impulsionnel ajoutés à une image simulée avec la variabilité précisée dans la légende. Les pixels détectés par le NFA sont présentés figure (b) et (c) : les vrais négatifs en vert, les vrais positifs en niveaux de gris, les faux négatifs en rouge et les faux positifs en jaune. Globalement, les faux négatifs et les faux positifs apparaissent très peu. . . . .	176
10.9. Erreurs (total) . . . . .	181
10.10 Non détections . . . . .	181
10.11 Non détections . . . . .	181
10.12 Erreurs (total) . . . . .	181
10.13 Non détections . . . . .	181
10.14 Non détections . . . . .	181
10.15 Erreurs (total) . . . . .	181
10.16 Non détections . . . . .	181
10.17 Non détections . . . . .	181
10.18 Erreurs de détection (% du nombre de pixels dans l'image) obtenues en fonction du taux d'occupation des changements dans un pixel <b>BR</b> . De gauche à droite, médiane et percentiles du total des erreurs obtenues, des faux positifs et des faux négatifs. Performances obtenues pour des images simulées avec 20% de pixels de changements et, de haut en bas, des rapports de résolution de 4, 16 et 32. . . . .	181
11.1. Bilan des mesures de terrain pour l'année agricole 2000 – 2001. . . . .	186
11.2. Images <b>HR</b> de réflectance acquises à 8 dates différentes, réparties entre le 14 novembre 2000 et le 31 juillet 2001. Les images acquises dans le canal 2 du capteur <b>SPOT/HRV</b> sont présentées en haut (TC2) et les images acquises dans le canal 3 sont présentées en bas (TC3). . . . .	187
11.3. Images <b>HR</b> de fraction de couvert acquises à 8 dates différentes, réparties entre le 14 novembre 2000 et le 31 juillet 2001. . . . .	188
12.1. A partir de la segmentation en 100 régions (a) (obtenue par le module segct du logiciel Megawave 2 pour deux images <b>HR</b> de fraction de couvert) et de la séquence d'images de fraction de couvert (b) de même résolution, l'algorithme non-supervisé permet d'obtenir la labellisation (c) pour 5 labels (labellisation de référence). . . . .	192
12.2. Labellisation obtenue pour 5 labels avec l'approche non-supervisée à partir de la segmentation (figure 12.1 (a)) et de la séquence <b>BR</b> de 8 images (a) de fraction de couvert (facteur de résolution $16 \times 16$ ). La carte des erreurs (c) montre (en blanc) environ 3.4% d'erreurs de labellisation (pourcentage de pixels mal labellisés) par rapport à la labellisation de référence (figure 12.1 (c)). . . . .	193

12.3. Labellisation obtenue pour 5 labels avec l'approche non-supervisée à partir de la segmentation (figure 12.1 (a)) et de la séquence <b>BR</b> de 8 images (b) de NDVI (facteur de résolution $16 \times 16$ ). La carte des erreurs (c) montre (en blanc) 3.6% d'erreurs de labellisation (pourcentage de pixels mal labellisés) par rapport à la labellisation de référence (figure 12.1 (c)). . . . .	193
12.4. Evolution temporelle des moyennes de fraction de couvert estimées par l'algorithme non-supervisé à partir de la séquence <b>HR</b> et de la séquence <b>BR</b> (moyennée par blocs de taille $16^2$ ). Les estimations réalisées à partir de données <b>BR</b> sont très proches de celles réalisées à partir de données <b>HR</b> . . . . .	194
12.5. Classification réalisée à partir de mesures sur le terrain (F. Oro, INRA). . . . .	194
12.6. Labellisations obtenues en mode non-supervisé à partir de la segmentation (a) contenant 100 régions et de la séquence <b>BR</b> de fraction de couvert (b) pour 4 à 13 labels (images (c) à (l)). . . . .	196
12.7. Labellisations obtenues en mode non-supervisé à partir de la segmentation (a) contenant 200 régions et de la séquence <b>BR</b> de fraction de couvert (b) pour 4 à 13 labels (images (c) à (l)). . . . .	197
12.8. Image <b>HR</b> de fraction de couvert du 14 novembre 2000 (image (a)) et classification (image (b)) réalisée par la méthode des K-moyennes pour 5 labels, à partir de l'image (a). . . . .	199
12.9. Validation de la classification 12.8 (b) par l'image 12.8 (a) dégradée d'un facteur $N = 5$ à $N = 50$ (figures (a) à (f)) avec la méthode de détection <i>a contrario</i> . Les pixels détectés comme non-cohérents avec la classification sont présentés en rouge, en surimpression sur l'image de fraction de couvert dégradée. L'évolution de la valeur du $NFA$ (en échelle $\log_{10}$ ) obtenue et le pourcentage du domaine validé sont tracés en fonction du facteur de résolution $N$ (figures (g) et (h)). Le domaine est bien validé par la méthode pour 95.5 à 96.6% des pixels de l'image. Remarquons que le domaine validé est stable en fonction du rapport de résolution $N$ . . . . .	200
12.10 Résultats obtenus avec la version non-vectorielle, à partir de la séquence de 8 images et des labellisations contenant de 1 à 20 labels (figure 12.6). Les valeurs du résidu cumulé, du $NFA$ (en échelle log) et du pourcentage du domaine validé sont respectivement tracées en fonction du nombre de labels contenus dans la labellisation. Chaque courbe représente les résultats obtenus pour 50 000, 500 000 et 5 000 000 itérations. Les résultats obtenus avec la version non-normalisée (modèle naïf : $\sigma^2 = \text{variance de la séquence}$ ) sont présentés à gauche et avec la version normalisée (images normalisées et $\sigma^2 = 1$ ) à droite. Les deux versions valident les labellisations contenant de 2 à 20 labels tant que le nombre d'itérations est suffisant (au-delà de 10 labels, l'augmentation des résidus montre l'absence de convergence). La labellisation pour un seul label est rejetée, par construction, par la version normalisée. . . . .	203

- 12.11 Résultats obtenus avec la version vectorielle, à partir de la séquence de huit images et des labellisations contenant de 1 à 20 labels (figure 12.6). Les valeurs du résidu cumulé, du  $NFA$  (en échelle log) et du pourcentage du domaine validé sont respectivement tracées en fonction du nombre de labels contenus dans la labellisation. Chaque courbe représente les résultats obtenus en fonction du nombre d'itérations considéré, suggérant que la convergence a été atteinte pour les labellisations contenant moins de 14 labels. Les résultats obtenus avec la version normalisée sont présentés à gauche et avec la version non-normalisée à droite. Les deux versions permettent de valider les labellisations contenant de 2 à 20 labels et marquent un minimum pour 12 labels. En revanche, seule la version normalisée permet, par construction, de rejeter la labellisation ne contenant qu'un seul label. . . . . 204
- 12.12 Résultats obtenus avec l'algorithme vectoriel normalisé pour les  $T$  premières dates de la séquence (figure 12.6 (b)) avec la labellisation (figure 12.6 (a)) réalisée pour les 8 images de la séquence. Les valeurs du résidu cumulé, du  $NFA$  (en échelle log) et du pourcentage du domaine validé sont respectivement tracées en fonction du nombre de labels contenu dans la labellisation. Chaque courbe représente les résultats obtenus en fonction de la séquence d'images considérée ( $T$  premières dates). La première image de la séquence ne permet pas, à elle seule, de valider les classifications. En revanche, les séquences des 2 à 8 premières images permettent de valider les classifications contenant de 2 à 15 labels (au-delà, l'augmentation des résidus montre que la convergence n'a pas été atteinte). Dans tous les cas, le  $NFA$  marque un minimum pour 12 labels. . . . . 206
- 12.13  $NFA$  minimal (en échelle log) obtenu pour des labellisations contenant de 5 à 20 labels et une image **BR** simulée à partir de 10 labels. Le minimum est bien obtenu pour 10 labels. . . . . 208
- 12.14 Evolution du  $NFA_2$  en fonction du nombre de labels. Résultats obtenus pour la séquence de huit images de fraction de couvert (figure 12.2 (a)) et les labellisations contenant 1 à 20 labels présentées figure 12.6 (en rouge) et 12.7 (en vert) (réalisées respectivement à partir de segmentations contenant 100 et 200 régions). La valeur minimale est atteinte pour 12 labels dans le cas des labellisations 12.6 ( $S = 100$ ), indiquant que la labellisation la plus pertinente pour la séquence d'images est celle réalisée avec 12 labels. Dans le cas des labellisations présentées figure 12.7 ( $S = 200$ ), un minimum est marqué pour 11 labels mais aussi pour 7 labels (valeur très proche). La pertinence des labellisations réalisées à partir de 200 régions est cependant plus faible que celle des labellisations réalisées à partir de 100 régions. 209



13.1. Résultats obtenus pour des changements introduits dans la classification par tirage aléatoire de 3, 4 ou 5 regions, et d'un nouveau label pour les régions tirées. Le label est tiré entre 1 et $L$ . Les changements introduits dans la labellisation sont représentés en jaune lorsqu'ils sont détectés et en vert sinon. Les pixels détectés qui ne correspondent pas à un changement sont représentés en rose s'ils étaient déjà détectés avant l'introduction de changements ( <i>cf.</i> (c)) et en rouge sinon. Les changements introduits sont globalement bien détectés, même lorsqu'ils concernent une faible portion d'un pixel <b>BR</b> . Dans l'exemple figure (c), un pixel de changement n'a pas été détecté (en haut à droite, en vert). . . . .	212
13.2. Détection de changements introduits sur l'image <b>BR</b> (b) correspondant à la classification (a) : les changements détectés sont représentés en rouge dans la figure (c) et la trace des changements introduits est représentée en blanc sur la même image. Les changements ont été introduits sur 5.46% des pixels de l'image <b>BR</b> et 96.2% ont été invalidés ( <i>cf.</i> figure 12.9 (c)). Les pixels détectés représentent 89.3% des pixels de l'image, valeur très proche des 90.7% attendus. . . . .	213
13.3. La labellisation (b) est simulée à partir de (a) en tant que labellisation de l'année précédente. L'image (c) présente en vert les différences entre les deux labellisations. La valeur du <i>NFA</i> et du pourcentage du domaine validé sont tracés en fonction du nombre de dates considéré. Le <i>NFA</i> permet, dès que 3 dates sont observées, de détecter les pixels qui ne sont pas cohérents avec la labellisation (en rouge, images (f) à (m)). . . . .	215
13.4. Simulation d'un changement sur les 4 dernières images d'une séquence de 8 acquisitions. La labellisation (a) (6 labels) correspond uniquement aux 4 premières dates. Les images (b) et (c) présentent, respectivement, l'image <b>BR</b> de la date 5 contenant un changement et l'image de différence faisant apparaître le changement introduit en niveaux de gris. Dans un contexte monotemporel, les images (d) à (g) représentent en rouge les pixels détectés en comparant la labellisation (a), successivement, aux images <b>BR</b> des dates 5, 6, 7 et 8 avant l'introduction de changements. Les images (h) à (k) représentent en rouge les pixels détectés à partir des images contenant les changements. Les pixels de changement (suffisamment marqués, <i>cf.</i> (c)) sont bien détectés à partir des dates 5 et 6. Ensuite, de nombreuses détections sont observées, que des changements aient été introduits ou non (figures (f), (g), (j) et (k)). Ces résultats semblent plutôt raisonnables puisque la labellisation considérée a été réalisée uniquement avec les 4 premières dates (d'hiver), une mise à jour est alors vraisemblablement nécessaire. . . . .	217

**Résumé :** Dans cette thèse, nous nous intéressons à l'analyse et au suivi temporel des surfaces continentales à partir de séquences d'images satellitaires. L'exploitation de données de différentes résolutions est alors cruciale pour bénéficier à la fois d'une bonne discrimination et d'une bonne localisation des objets d'intérêt. Dans ce contexte, nous proposons deux approches probabilistes pour la classification et la détection de changements capables d'accéder à une information sous-pixelique, avec très peu d'information *a priori*. La première repose sur la définition d'une fonction d'énergie dans un cadre bayésien. Étant donné un nombre de classes, elle permet d'estimer la classification de manière non-supervisée en tant que minimum de cette fonction d'énergie, à travers un algorithme de recuit simulé. La seconde repose sur un modèle de détection *a contrario* couplé à un algorithme stochastique d'échantillonnage aléatoire. Elle permet de détecter automatiquement les pixels de l'image qui représentent le plus vraisemblablement des changements. Une analyse théorique et expérimentale des méthodes proposées a permis d'en cerner les limites et, en particulier, de montrer leur capacité à traiter de forts rapports de résolution. Des cas réels d'applications sont présentés sur une scène agricole de la Plaine du Danube (base de données ADAM).

**Mots-clés :** Traitement d'images, détection de changements, classification, détection a-contrario, modèles bayésiens, imagerie satellitaire, séquences temporelles, analyse sous-pixelique, désagrégation de données basse résolution, algorithmes stochastiques, recuit simulé.

---

**Abstract :** This thesis focuses on the land cover analysis and monitoring from remote sensing time series. The use of data with different resolution is critical for both a good discrimination and a good localization of the objects of interest. In this context, we propose two approaches for sub-pixelic classification and change detection, using very few *a priori* information. The first one is based on the definition of an energy function in a Bayesian framework. Given a number of classes, it enables an unsupervised estimation of the classification as a minimum of this energy function, through a simulated annealing algorithm. The second one is based on an a-contrario detection model with a stochastic algorithm that automatically selects the image subdomain representing the most likely changes. A theoretical and experimental analysis of the proposed approaches enabled to estimate their limitations and, in particular, to show their capability to deal with high resolution ratios. Actual applications are presented in the case of an agricultural scene of the Danubian plain (ADAM database).

**Keywords :** Image processing, change detection, classification, a-contrario detection, Bayesian models, remote sensing, time series, sub-pixelic analysis, disaggregation of coarse resolution data, stochastic algorithm, simulated annealing.